ELSEVIER

Knowledge-Based Systems 13 (2000) 93–99

Knowledge-Based SYSTEMS

www.elsevier.com/locate/knosys

Multiple algorithms for fraud detection

R. Wheeler*, S. Aitken

Artificial Intelligence Applications Institute, The University of Edinburgh, 80 South Bridge, Edinburgh EH1 1HN, Scotland, UK

Abstract

This paper describes an application of Case-Based Reasoning to the problem of reducing the number of final-line fraud investigations in the credit approval process. The performance of a suite of algorithms, which are applied in combination to determine a diagnosis from a set of retrieved cases, is reported. An adaptive diagnosis algorithm combining several neighbourhood-based and probabilistic algorithms was found to have the best performance, and these results indicate that an adaptive solution can provide fraud filtering and case ordering functions for reducing the number of final-line fraud investigations necessary.

Keywords: Fraud detection; Case-based reasoning; Adaptive algorithms

1. Introduction

Artificial intelligence techniques have been successfully applied to credit card fraud detection and credit scoring, and the field of AI as applied to the financial domain is both well-developed and well documented. As an emerging methodology, case-based reasoning (CBR) is now making a significant contribution to the task of fraud detection. CBR systems are able to learn from sample patterns of credit card use to classify new cases, and this approach also has the promise of being able to adapt to new patterns of fraud as they emerge. At the forefront of research in this field is the application of adaptive and hybrid learning systems to problems which previously were considered too dynamic, chaotic, or complex to accurately model and predict.

As applied to the financial domain, CBR systems have a number of advantages over other AI techniques as they:

- provide meaningful confidence and system accuracy measures;
- require little or no direct expert knowledge acquisition;
- are easily updated and maintained;
- articulate the reasoning behind the decision making clearly;
- are flexible and robust to missing or noisy data;
- may take into account the cost effectiveness ratio of investigating false positives and advise accordingly; and
- are easily integrated into varying database standards.

And the addition of adaptive CBR components may allow the system to:

- optimise the accuracy of classification by dynamically adjusting and updating weighting structures;
- use multiple algorithms to enhance final diagnostic accuracy; and
- better differentiate between types of irregularities and develop a diagnostically significant sense of abnormality which aids in the first-time detection of new irregularity types.

In this paper we describe the background of a complex fraud-finding task, and then describe the development of an adaptive proof-of-concept CBR system, which is able to achieve very encouraging results on large, noisy real-world test sets. In specific, this paper addresses the problem of making a diagnostic decision given a set of near-matching cases. Finally, the results of the investigation are summarised and considered in the light of other work in the field.

2. Background

At the request of one of the UK's most successful fraud detection system software providers, AIAI undertook an investigation into methods of applying new AI technologies to increase the accuracy of the already highly advanced systems presently in use. While the firm's software presently reduces the number of necessary fraud investigations by several orders of magnitude, our investigation showed that utilising adaptive algorithms and fuzzy logic

* Corresponding author. Tel.: +44-1316502732; fax: +44-1316506513.
E-mail address: richardw@aiai.ed.ac.uk (R. Wheeler).

0950-7051/00/$ - see front matter
PII: S0950-7051(00)00050-2

results in a significant diagnostic improvement on the most difficult sub-section of cases.

The focus of our investigation was to reduce the number of applications referred for expert investigation after the existing detection systems had been utilised. These well-proven systems take advantage of many person years of expert knowledge elicitation and encoding, and are able to reduce the initial volume of applications by roughly 2500 times (400 referred from every million analysed). It was on a database consisting solely of these hardest cases that the CBR system attempted to make diagnostically significant decisions.

The source data comprised pairs of database records, the first of which was tagged as the application, the second as the evidence found by previous analysis suggesting fraud. Two files of data were provided: one set of nearly 2000 application-match pairs which were initially flagged as fraud but later cleared (the non-fraud set), and a second set of 175 application-match pairs which were judged to be fraudulent.

The application records consisted of applicant data (personal code, name of the lender, amount of loan, name and address of applicant, etc.), and employer data (type of business, time employed, etc.). Evidence provided followed a similar format, and the final test sets consisted of 584 cases of non-fraud and 96 cases of fraud. Each case consisted of an application and one or more evidence records both of which contained application and employer data.

Pre-processing was kept to a minimum, excluding only those fields which might be construed as being false indicators, such as database tags generated by the company's selection process. All other fields remained in their original state, and omissions formed a high percentage of total information encoded.

In order to capture more general patterns in application-match pairs *within* a case, the type of match that existed between fields was introduced into the case description. A small number of terms were defined to describe these matches, and this information was added into the cases after parsing. These general descriptions of matches were given simple descriptive labels: exact-match, near-match, dissimilar and added as a third component to each application-match pair. As such, the additional information was intended to act as a general fuzzy classifier of match fitness. The conjecture was that there are patterns of values for match types that might be exploited by an adaptive system. Similarity measures were assessed for all field types: strings, dates, addresses, numerical values, etc., and it was these final three-part sets: application, evidence, fuzzy match descriptor, that were presented to the proof-of-concept system for analysis. After all pre-processing, each case was described by 128 attributes.

3. Approach

Statistical investigations of the test sets suggested that the nature of the problem is inherently non-linear, noisy, contradictory, and not addressable using a simple similarity matrix and CBR decision system. This is unsurprising as the test sets were composed of the most difficult and intractable subset of the credit approval data, and as such did not cluster into identifiable fraud/non-fraud regions. However, highly localised phenomena and patterns appeared fairly common, suggesting that a hybrid or adaptive system within a CBR methodological structure might be able to focus upon and effectively exploit these characteristics.

The proof-of-concept system design has two essential decision-making components familiar to all CBR frameworks: retrieval and diagnosis. Retrieval utilises a weight matrix and nearest neighbour algorithm, while diagnosis utilises a suite of algorithms, which analyse the data recalled by the retrieval mechanism as being significant. A learning mechanism was also implemented in the proof-of-concept system.

In this section we present the investigation of weighting matrix approaches, nearest neighbour strategies employed, and multi-algorithm final analysis, which is the focus of this work.

3.1. Weighting matrix approach

CBR systems function by defining a set of features within a data or case base, and then generating a similarity score that represents the relationship between a previously seen case and the test case. Generally, this comparison is flat, that is, each field matching according to predefined operators (such as exact and fuzzy matching) adds one point to the total similarity score of the comparison. Of course, not all fields (or features) within a database are equally meaningful in divining a classification or sound decision, so a weighting matrix is often employed—a method by which a single feature's importance may be raised or lowered, giving certain features more diagnostic significance than others.

The first set of experiments performed were to experimentally assess the effect on total diagnostic accuracy of the raising and lowering of individual field weights. This was performed with the AIAI CBR Shell System[1] which supports the automatic polling of fields for sensitivity to goal finding and the stochastic hill-climbing of ever-fitter combinations of field weights. Disappointingly, these investigations only demonstrated that any simple relationships between field values and fraud occurrence had already been exploited by the rule-based filtering that had been applied to the data prior to our analysis of it. In consequence, a flat weighting structure was used in all subsequent testing.

3.2. Case retrieval

Nearest neighbour matching is common to many CBR systems. Again using the basic exploratory facilities of

[1] See: http://www.aiai.ed.ac.uk/~richardw/cbrshell.html.