Outlier Detection Using Replicator Neural Networks

Simon Hawkins, Hongxing He, Graham Williams, and Rohan Baxter

CSIRO Mathematical and Information Sciences
GPO Box 664, Canberra ACT 2601, Australia
Firstname.Lastname@csiro.au

Abstract. We consider the problem of finding outliers in large multivariate databases. Outlier detection can be applied during the data cleansing process of data mining to identify problems with the data itself, and to fraud detection where groups of outliers are often of particular interest. We use replicator neural networks (RNNs) to provide a measure of the outlyingness of data records. The performance of the RNNs is assessed using a ranked score measure. The effectiveness of the RNNs for outlier detection is demonstrated on two publicly available databases.

1 Introduction

Outlier detection algorithms have application in several tasks within data mining. Data cleansing requires that aberrant data items be identified and dealt with appropriately. For example, outliers are removed or considered separately in regression modelling to improve accuracy. Detected outliers are candidates for aberrant data. In many applications outliers are more interesting than inliers. Fraud detection is a classic example where attention focuses on the outliers because these are more likely to represent cases of fraud. Fraud detection in insurance, banking and telecommunications are major application areas for data mining. Detected outliers can indicate individuals or groups of customers that have behaviour outside the range of what is considered 'normal' [8,6,21].

Studies from the field of statistics have typically considered outliers to be residuals or deviations from a regression or density model of the data:

> An outlier is an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism [9].

In this paper we employ multi-layer perceptron neural networks with three hidden layers, and the same number of output neurons and input neurons, to model the data. These neural networks are known as replicator neural networks (RNNs). In the RNN model the input variables are *also* the output variables so that the RNN forms an implicit, compressed model of the data during training. A measure of outlyingness of individuals is then developed as the reconstruction error of individual data points. The RNN approach has linear analogues in Principal Components Analysis [10].

Y. Kambayashi, W. Winiwarter, M. Arikawa (Eds.): DaWaK 2002, LNCS 2454, pp. 170–180, 2002.

The insight exploited in this paper is that the trained neural network will reconstruct some small number of individuals poorly and these can be considered as outliers. We measure outlyingness by ranking data according to the magnitude of the reconstruction error. This compares to SmartSifter [22] which similarly builds models to identify outliers but scores the individuals depending on the degree to which they perturb the model.

Following [22], [4] and [17] when dealing with large databases, we consider it more meaningful to assign each datum an *outlyingness score*. The continuous score reflects the fuzzy nature of outlyingness and also allows the investigation of outliers to be automatically prioritised for analysis.

2 Related Work

We classify outlier detection methods as either *distribution-based* or *distance-based*. However, a probabilistic interpretation can often be placed on the distance-based approaches and so the two categories can overlap. Other classifications of outlier detection methods are based on whether the method provides an outlyingness score or a binary predicate (which may also be based on a score), or whether the method measures outlyingness from the *bulk* (i.e., a convex hull) of the data, or from a regression surface. Distribution-based methods include mixture models such as SmartSifter [22]. Individuals are scored according to the degree to which they perturb the currently learnt model. Distance-based methods use distance metrics such as Mahalanobis distance [2,15] or Euclidean distance [11,13,12]. A prominent and useful technique for detecting outliers is to use a clustering algorithm, such as CURE or BIRCH, and then designate data occurring in very small clusters, or data distant from existing clusters as outliers [16,21,7,23,14]. Visualisation methods [3], based on grand tour projections, can also be considered distance-based since the distance between points is projected onto a 2-dimensional plane. Visualisations using immersive virtual environments [20] similarly explore the space for outliers allowing users to identify and view outliers in multiple dimensions. Despite the obvious issues of subjectiveness and scaling, visualisation techniques are very useful in outlier detection. Readily available visualisation tools such as *xgobi* [18] provide an effective, efficient, and interactive initial exploration of outliers in data (or necessarily a sample of the data). We are aware of only one other previous neural network method approach to detecting outliers [19]. Sykacek's neural network approach is to use a multi-layer perceptron (MLP) as a regression model and to then treat outliers as data with residuals outside the error bars.

3 Replicator Neural Network Outlier Detection

Although several applications in image and speech processing have used the Replicator Neural Network for its data compression capabilities [1,10], we believe the current study is the first to propose its use as a outlier detection tool.