International Journal of Approximate Reasoning 85 (2017) 68–78

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Cost-sensitive sequential three-way decision modeling using a deep neural network [☆]

Huaxiong Li [a,*,1], Libo Zhang [a,1], Xianzhong Zhou [a], Bing Huang [b,*]

[a] Department of Control and Systems Engineering, School of Management and Engineering, Nanjing University, Nanjing, 210093, China
[b] School of Technology, Nanjing Audit University, Nanjing, Jiangsu, 211815, PR China

ARTICLE INFO

Article history:
Received 29 June 2016
Received in revised form 18 March 2017
Accepted 20 March 2017
Available online 23 March 2017

Keywords:
Cost-sensitive learning
Deep neural network
Granular computing
Three-way decision

ABSTRACT

Three-way decision (3WD) models have been widely investigated in the fields of approximate reasoning and decision making. Recently, sequential 3WD models have attracted increasing interest, especially for image data analysis. It is essential to select an appropriate feature extraction and granulation method for sequential 3WD-based image data analysis. Among the existing feature extraction methods, deep neural networks (DNNs) have been considered widely due to their powerful capacity for representation. However, several important problems affect the application of DNN-based feature extraction methods to sequential 3WD. First, it takes a long time for a DNN to obtain an optimal feature representation. Second, most DNN algorithms are cost-blind methods and they assume that the costs of all misclassifications are the same, which is not the case in real-world scenarios. Third, DNN algorithms are two-way decision models and they cannot provide boundary decisions if sufficient information is not available. To address these problems, we propose a DNN-based sequential granular feature extraction method, which sequentially extracts a hierarchical granular structure from the input images. Based on the sequential multi-level granular features, a cost-sensitive sequential 3WD strategy is presented that considers the misclassification cost and test cost in different decision phases. Our experimental analysis validated the effectiveness of the proposed sequential DNN-based feature extraction method for 3WD.

1. Introduction

In traditional two-way decision models, there are only two choices when making a decision: a positive decision or negative decision. An incorrect decision is always associated with a high cost, so a two-way decision model may lead to a high-cost loss due to the lack of suitable information for making a precise decision. To address this issue, Yao proposed a three-way decision model (3WD), where the boundary decision is considered to be the optional choice and it is added to the decision actions set [45]. In the past decade, 3WD has received much attention in the fields of uncertain reasoning and decision making, such as medical decision making [44], classification [30,53], uncertainty management [6,11,12], email spam filtering [14,59], clustering analysis and covering reduction analysis [51–53], incomplete data analysis [19,33,34,42],

☆ This paper is part of the virtual special issue on tri-partition, edited by Davide Ciucci and Yiyu Yao.

* Corresponding authors.
E-mail addresses: huaxiongli@nju.edu.cn (H. Li), zhanglb@smail.nju.edu.cn (L. Zhang), zhouxz@nju.edu.cn (X. Zhou), hbhuangbing@126.com (B. Huang).

[1] The first two authors made equivalent contributions and should be considered co-first authors.

http://dx.doi.org/10.1016/j.ijar.2017.03.008
0888-613X/

and knowledge discovery [3]. Traditional 3WD considers a static and one-step decision strategy. The costs of all the decision actions are computed based on the currently available information and the optimal decision is determined by minimizing the decision cost via Bayesian theory. In the case where sufficient information is available to make a decision, a low-cost and precise decision can be made using the currently available information. By contrast, in the case where the available information is insufficient, most decisions may fall into a boundary region because positive decision and negative decisions are both associated with high costs. In this case, it is necessary to acquire more information to further improve the precision of the decision. Thus, the decision cost decreases and newly available information is then employed to enhance the precision of the decision. This process continues until a satisfactory low-cost decision is reached in a sequential decision strategy [50, 39].

A sequential 3WD model was first proposed by Yao in 2011 [50]. Recently, a series of sequential 3WD models were developed based on a decision-theoretic rough sets framework [13,21,22,27,31], which provide effective methods for information with multiple levels of granularity and a hierarchical granular structure representation [48,43,25,24,38]. Li et al. presented a cost-sensitive sequential 3WD model for data classification and image recognition [20,22]. Clearly, the cost-sensitive sequential 3WD method is consistent with real-world applications in face recognition and identification [20]. If we consider a criminal investigation scenario, the facial images captured by a surveillance camera may be too blurred for use in identification. Thus, it is necessary to collect better quality facial images using a high definition surveillance camera. The results of criminal investigations based on a series of high-quality facial images will always yield a satisfactory low-cost decision. However, a high definition surveillance camera may be expensive, which increases the test cost. Thus, it is necessary to search for an appropriate sequential decision strategy to balance the decision and test costs, thereby achieving a low cost overall decision [20,56].

It is essential to find an appropriate feature representation and image granulation method for sequential 3WD-based facial image data analysis. Similarly, in the field of face recognition, the importance of the feature extraction method employed cannot be overstated. Thus, numerous feature extraction and representation methods have been developed [58,2]. Recently, deep neural networks (DNNs) have received much attention in the fields of feature extraction and image recognition because of their high capacity for representation and image recognition performance [2]. In general, a DNN is a multi-layer artificial neural network comprising input neurons, output neurons, and multi-layer hidden neurons. Hinton et al. proposed an effective DNN model based on a deep sparse auto-encoder (SAE) using a restricted Boltzmann machine [9]. It was demonstrated that deep SAE is a high-performance feature representation and learning method because it has the following two advantages: (1) DNN with multiple hidden layers has an excellent capacity for feature learning and the essential features extracted from the DNN can greatly facilitate various visualization and classification tasks [9]; (2) DNN incorporates a greedy learning algorithm, which can find a fairly good set of parameters within a relatively short training process, even in deep networks with a large number of parameters and many hidden layers [9]. The first DNN approach to achieve equivalent or even better performance than humans was proposed by Ciresan [4]. In recent years, DNN algorithms have been applied successfully to automatic image captioning [35], video analysis [60], and in many fields of artificial intelligence research [1].

Despite their optimal performance in various fields, some important issues with DNNs still need to be addressed in uncertain reasoning and complex problem solving. First, although DNNs employ a greedy level-wise search algorithm to find the objective parameters set, obtaining an optimal feature representation is still very time consuming, especially in deep networks with a large number of parameters and many hidden layers. In reality, a more important evaluation measure is the decision cost rather than the decision accuracy. In this case, it is not necessary to spend as much time on the training process to obtain a high accuracy decision result. A rapid boundary decision or a fast low-cost decision may be a satisfactory decision; otherwise, more information is required to obtain a low-cost decision. This decision process is highly consistent with the framework of sequential 3WD [20]. Second, to the best of our knowledge, most DNN algorithms are cost-blind methods, where they assume that the costs of all misclassifications are the same, but this assumption is not reasonable in real-world applications [16,54,57]. If we consider an entrance guard system as an example, it may be inconvenient if an office member is misrecognized as an imposter, but it would be far more disastrous if an imposter is misrecognized as an office member and enters the office [57]. Thus, it is necessary to incorporate cost information in DNNs, so cost-sensitive decisions and classification problems can be investigated using DNNs. Third, a DNN is an effective feature learning method, but the last layer neurons in DNN models are always connected to a classifier for classification and decision. Thus, the classification and decision accuracies are highly correlated with the connected classifiers. Most existing classifiers used in DNN models are two-way decision models and they cannot provide boundary decisions. Therefore, it is necessary to introduce boundary decisions into DNN classifiers if the currently available information is not sufficient and uncertain.

In order to address these problems, we propose a cost-sensitive sequential 3WD-based DNN algorithm for image data analysis and image recognition. Considering the different levels of available information that can be used for image recognition, the features extracted from different training loops in a DNN are represented as the multi-level granularity of images. At each level of granularity, an optimal decision is computed based on 3WD, which outputs a minimum cost decision, including positive, negative, and boundary decisions, according to the currently available information. Cost-sensitive 3WD is conducted at each level of granularity for the images extracted in each training loop of the DNN, where the misclassification and test costs are both considered. As the number of training loops increases, the misclassification cost decreases whereas the test cost increases. The sequential decision process terminates in the decision step when the minimal total cost is reached. This sequential decision process simulates a human decision-making process where an uncertain sequential decision process progresses from a rough granularity level to a precise granularity level.