Information Sciences 447 (2018) 36–51

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Self-organising fuzzy logic classifier

Xiaowei Gu[a], Plamen P. Angelov[a,b,*]

[a] School of Computing and Communications, Lancaster University, Lancaster LA1 4WA, UK
[b] Honorary Professor, Technical University, Sofia 1000, Bulgaria

ARTICLE INFO

Article history:
Received 22 September 2017
Revised 1 March 2018
Accepted 3 March 2018
Available online 6 March 2018

Keywords:
Classification
Fuzzy rule-based systems
Self-organising
Recursive

ABSTRACT

In this paper, we present a self-organising nonparametric fuzzy rule-based classifier. The proposed approach identifies prototypes from the observed data through an offline training process and uses them to build a 0-order AnYa type fuzzy rule-based system for classification. Once primed offline, it is able to continuously learn from the streaming data afterwards to follow the changing data pattern by updating the system structure and meta-parameters recursively. The meta-parameters of the proposed approach are derived from data directly. By changing the level of granularity, the proposed approach can make a trade-off between performance and computational efficiency, and, thus, the classifier is able to address a wide variety of problems with specific needs. The classifier also supports different types of distance measures. Numerical examples based on benchmark datasets demonstrate the high performance of the proposed approach and its ability of handling high-dimensional, complex, large-scale problems.

1. Introduction

Classification is one of the hotly studied problems in machine learning [24]. Till now, various classification algorithms have been successfully developed and widely used in different areas i.e. remote sensing [46,47], face recognition [10,25], handwritten digits recognition [13,21], etc.

Current classification approaches have different architectures. In general, considering their operating mechanisms, the existing approaches can be categorised into two major types: *1)* offline [13,14,27] and *2)* online [6,8,21,31,34,38,39,44]. The offline approaches are trained with static datasets and once the training process is finished, the classifiers stop learning and allow no further modification to their structure. The majority of the offline approaches were developed during the time that data was not considered to be in large-scale, streaming and dynamically evolving. Nowadays, as we are living in the era of the so-called "Big Data", these approaches become less applicable. There are two types of online classification approaches, namely, *1)* incremental [31,34,44] and *2)* evolving [6,8,21,38,39]. Online approaches can be of "one-pass" type, which means that they are able to consistently learn from newly arrived data samples and only store the key information in memory, meanwhile, discard all the processed training samples. The evolving approaches [6,8,21,38,39], as the more advanced branch of online approaches, further address the problem of changing data pattern in nonstationary environments by continuously evolving system structure and recursively updating meta-parameters. Compared with the other types, evolving approaches are more memory- and computation- efficient and, thus, are more frequently used in real-world applications. On the other hand, the performance of the online approaches, including the evolving ones, is sensitive to the order of data samples.

* Corresponding author at: School of Computing and Communications, Lancaster University, Lancaster LA1 4WA, UK.
E-mail addresses: x.gu3@lancaster.ac.uk (X. Gu), p.angelov@lancaster.ac.uk (P.P. Angelov).

https://doi.org/10.1016/j.ins.2018.03.004
0020-0255/

Very often in real situations, a part of the data is available in a static form, while the rest is observed sequentially in a streaming form. Offline approaches ignore the fact that the data pattern may change with more data available. However, it is also unnecessary for an approach to learn online from the very beginning of the data stream because initialising the system with the available static data in an offline manner can guarantee a more robust performance.

Furthermore, many existing approaches also rely heavily on *1) prior* assumptions, which usually impose models with parameters which depend on the data generation model, i.e. Gaussian distribution [29], and *2)* user inputs, which are defined based on *prior* knowledge of the problem, i.e. radius [15,21,50], learning rate [38]/decay rates [39], size of the network [13,23], etc. In real cases, such *prior* assumptions are often too strong to be held and user inputs are often hard to define due to the insufficient *prior* knowledge. In addition, in online scenarios, non-stationary data streams may also invalidate the *prior* assumptions and user inputs that were established at the initial stage.

In this paper, a new self-organising fuzzy logic (SOF) approach is proposed for classification. The SOF approach is grounded at the recently introduced Empirical Data Analytics (EDA) computational framework [4,5] and the autonomous data-driven clustering techniques [19]. The SOF classifier has two training stages, *1)* offline and *2)* online. During the offline stage, it learns from the static data to establish a stable 0-order AnYa type fuzzy rule-based (FRB) system [7]. During the online training stage, the FRB system identified through the offline training process will be updated subsequently with the streaming data to follow the possible *drifts* and/or *shifts* in the data pattern. The SOF classifier only keeps the key meta-parameters in memory and is of "one-pass" type during its online training stage; therefore, it is very suitable for large-scale streaming data processing.

Most importantly, the proposed SOF classifier is nonparametric in the sense that no parameters or models are imposed for the data generation model. Employing the EDA quantities as described in section 2.2, the SOF classifier is able to objectively disclose the ensemble properties and mutual distributions of the streaming data based on the empirical observations and all the meta-parameters of the classifier are directly derived from the data without any *prior* knowledge [4,5].

The proposed SOF classifier keeps the advantage of objectiveness of the data-driven approaches, and, at the same time, puts users "in the driving seat" by letting users to decide the level of granularity and the type of distance/dissimilarity measure for it. The idea of "granularity" is introduced and defined in [35,36,49]. It is well known that a problem can be approached at different levels of specificity (detail) depending on the complexity of the original problem, available computing resources, and particular needs [49]. The level of granularity in the proposed approach is aligned with this concept. However, it has to be stressed that there is no requirement for *prior* knowledge to decide the level of granularity and it can be given merely based on the preferences of the users. Higher level of granularity leads to a classifier with fine details, and at the same time, results in a risk of overfitting. A lower level of granularity, instead, gives users a classifier trained coarsely but with higher computational efficiency, generalisation and less memory requirement. The SOF classifier is always guaranteed to be meaningful due to its data-driven nature. The choice of the type of distance/dissimilarity measure further gives more freedom to the users and also makes the proposed SOF approach highly adaptive to various applications, e.g. natural language processing. In addition, the SOF classifier can also provide the default level of granularity and distance measure option for the less experienced users.

The remainder of this paper is organised as follows. The theoretical basis of the SOF classifier is summarised in Section 2. Section 3 describes the offline training, online training and validation processes of the proposed approach. Section 4 presents how the level of granularity can influence the performance and efficiency of the SOF classifier. Numerical examples serving as a proof of concept are given in Section 5, discussions on the convergence and local optimality of the proposed approach are also provided in the same section. Section 6 concludes this paper and gives the direction for future works.

2. Theoretical basis

In this section, the theoretical basis of the self-organising fuzzy logic (SOF) classifier will be briefly summarised.

2.1. 0-order AnYa fuzzy rule-based systems

AnYa type FRB system was introduced in [7] as an alternative approach to the widely used FRB systems of Takagi-Sugeno [45] or Mamdani [30] types. Comparing with the two predecessors, the antecedent (IF) part of AnYa type fuzzy rules is simplified to a more compact, objective and nonparametric vector form without the need of defining *ad hoc* membership functions. A 0-order AnYa type fuzzy rule has the following form:

$$IF\ (\boldsymbol{x} \sim \boldsymbol{p}_1)\ OR\ (\boldsymbol{x} \sim \boldsymbol{p}_2)\ OR \ldots OR\ (\boldsymbol{x} \sim \boldsymbol{p}_N) \quad THEN\ (class) \tag{1}$$

where $\boldsymbol{x}$ is the input vector; "$\sim$" denotes similarity, which can also be seen as a fuzzy degree of satisfaction/membership [7]; $\boldsymbol{p}_i$ $(i = 1, 2, \ldots, N)$ is the ith prototype of the class; N is the number of prototypes identified from the data samples of this class. For a specific data sample, its label can be decided following different strategies, i.e. "winner-takes-all", "few-winners-take-all", "fuzzily weighted average", etc. In this paper, we use the first one, and the details are given in Section 3.3.

2.2. Empirical data analytics operators

As stated in Section 1, the SOF classifier employs the nonparametric EDA quantities for objectively disclosing the ensemble properties and mutual distribution of the data. In this subsection, three EDA quantities, *1) cumulative proximity, 2)*