Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 35 (2014) 358 – 367

Procedia Computer Science

18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems - KES2014

Parallel learning and classification for rules based on formal concepts

Nida Meddouri[a,*], Hela Khoufi[a], Mondher Maddouri[a,b]

[a]*Laboratoire d'Informatique, Programmation, Algorithmique et Heuristique - LIPAH*
Faculté des Sciences mathématiques, physiques et naturelles de Tunis - FST
Université d'El Manar
Campus universitaire El Manar, 1060, Tunis, Tunisie.
[b]*Department of Computer Sciences, Community College of Hinakya;*
Taibah University, Medina, Kingdom of Saudi Arabia.

Abstract

Supervised classification is a spot/task of *data mining* which consist on building a classifier from a set of instances labeled with their class (*learning step*) and then predicting the class of new instances with a classifier (*classification step*). In supervised classification, several approaches were proposed such as: *Induction of Decision Tree* and *Formal Concept Analysis*. The learning of formal concepts is generally based on the mathematical structure of *Galois lattice* (or *concept lattice*). The complexity of *Galois lattice* generation limits the application fields of these systems. In this paper, we discuss about supervised classification based on *Formal Concept Analysis* and we present methods based on *concept lattice* or *sub lattice*. We propose a new approach that builds only a part of the lattice, including the best concepts (i.e pertinent concepts). These concepts are used as classifiers in parallel combination using voting rule. The proposed method is based on *Dagging of Nominal Classifier*. Experimental results are given to prove the interest of the proposed method.

Peer-review under responsibility of KES International.
Keywords: Formal Concept, Ensemble method, Dagging, Classification Rules, Machine Learning, Data Mining.

1. Introduction

Formal Concept Analysis is a formalization of the philosophical notion of concept defined as a couple of extension and comprehension. The comprehension (called also intention) makes reference to the necessary and sufficient attributes which characterizes this concept. The extension is a set of instances which made it possible to find out the concept [1].

* Corresponding author. Tel.: +216-20-422-463.
E-mail address: nida.meddouri@gmail.com

1877-0509
Peer-review under responsibility of KES International.
doi:10.1016/j.procs.2014.08.116

The classification approach based on *Formal Concept Analysis* is a symbolic approach allowing the extraction of correlations, reasons and rules according to the concepts discovered from data. Supervised classification is a process made up of two steps. In the learning step, we organize the information extracted from a group of objects in the form of a lattice. In the classification step, we determine the class of new objects, based on the extracted concepts. Many learning methods based on *Formal Concept Analysis* were proposed, such as: `GRAND`[2], `CLNN&CLNB`[3], `IPR`[4], `NAVIGALA`[5], `CITREC`[6] and more recent `BFC`[7] and `BNC`[8]. Unfortunately, systems based on *Formal Concept Analysis* encountered some problems such as exponential complexity (in the worth case), high error rate and over-fitting[8,9].

In this last decade, a great number of researches in machine learning have been concerned with the ensemble methods of classifiers that allow the improvement of a single learner performance (generally a weak learner) by the voting techniques[10]. In the area of supervised learning, several ensemble methods have been appeared[11] such as *Boosting* and *Bagging* which improve the performance of combined classifiers sets. The two principal reasons for this success are probably the simplicity of implementation and the recent theorems relative to the boundaries, the margins, or to the convergence[10,12,13]. Generally, the ensemble methods are based on sequential or parallel learning (*Bagging*). The difference between them derives from how to select data for learning.

In sequential learning such as *Boosting*, all the data are considered in each learning step and the weights are assigned to learning instances. However, it was proved that this method is not interesting and no sufficient for a more efficient classifier as *Decision Tree*[8]. In parallel learning, such as *Bagging*, the training data are drawn randomly with replacement from the original data set, such a training set is called a *Bootstrap*. The well known method which is based on parallel learning is *Dagging* (***D**isjoint samples **agg**regat**ing***), it creates a number of disjoint groups and stratified data from the original learning data set, each one is considered as a subset of learning. The weak learner is built on this learning sets. The predictions are then obtained by combining the classifier outputs by majority voting[14]. *Dagging* has shown its importance in recent work. Then, we propose to use this technique, in this work, to study the classifier ensembles based on formal concepts, since, no study has focused on the formal concepts in the context of parallel learning.

In section 2, we present a state of the art on *Formal Concept Analysis* and several methods used which are based on lattice concept and sub-lattice of concepts. In section 3, we propose a new method exploiting the advantages of the *Dagging* to generate and combine in parallel way weak concept learners[15 16]. From the section 4, a comparative experimental study is presented to evaluate the performance of parallel classifier ensembles according to certain criteria such as the number, variety and the type of classifiers. A comparative experiment is also presented to show the importance of parallel learning using stratified sampling, compared to sequential learning using random sampling.

2. Formal concept analysis and classification

2.1. Definition

A formal context is a triplet $< \mathcal{O}, \mathcal{P}, \mathcal{R} >$, where $\mathcal{O} = \{o_1, o_2, ..., o_n\}$ is a finite set of n instances, $\mathcal{P} = \{p_1, p_2, ..., p_m\}$ a finite set of m properties (binary attributes) and $\mathcal{R}$ is a binary relation defined between $\mathcal{O}$ and $\mathcal{P}$. The notation $(o_i, p_j) \in \mathcal{R}$ or $\mathcal{R}(o_i, p_j) = 1$ means that the instance o_i verifies the property p_j in relation $\mathcal{R}$[1]. The context (see Table 1 and 2)[1] is often represented by a cross-table or a binary-table.

Let $A \subseteq \mathcal{O}$ and $B \subseteq \mathcal{P}$ be two finite sets. For both sets A and B, operators $\varphi(A)$ and $\delta(B)$ are defined as[1]:

- $\varphi(A) = \{p \mid \forall o,\ o \in A \ and \ (o, p) \in \mathcal{R}\}$.
- $\delta(B) = \{o \mid \forall p,\ p \in B \ and \ (o, p) \in \mathcal{R}\}$.

Operator φ defines the properties shared by all elements of A. Operator δ defines instances which share the same properties included in set B. Operators φ and δ define a `Galois` connexion between sets $\mathcal{O}$ and $\mathcal{P}$[1]. The closure operators are $A" = \varphi \circ \delta(A)$ and $B" = \delta \circ \varphi(B)$. Finally, the closed sets A and B are defined by $A = \varphi \circ \delta(A)$ and $B = \delta \circ \varphi(B)$.

[1] The data sets is selected from `UCI Machine Learning Repository`[17]