Evidential Learning Classifier System

Chedi Abdelkarim, Lilia Rejeb, Lamjed Ben Said, Maha Elarbi
chediabd@gmail.com,lilia.rejeb@gmail.com,bensaid_lamjed@yahoo.fr,arbi.maha@yahoo.com
Université de Tunis, ISGT, LR11ES03 SMART Lab
41, Avenue de la Liberté, Bouchoucha
Bardo, Tunisia 2000

ABSTRACT

During the last decades, Learning Classifier Systems have known many advancements that were highlighting their potential to resolve complex problems. Despite the advantages offered by these algorithms, it is important to tackle other aspects such as the uncertainty to improve their performance. In this paper, we present a new Learning Classifier System (LCS) that deals with uncertainty in the class selection in particular imprecision. Our idea is to integrate the Belief function theory in the sUpervised Classifier System (UCS) for classification purpose. The new approach proved to be efficient to resolve several classification problems.

CCS CONCEPTS

•**Data Mining → Uncertainty;**

KEYWORDS

Learning Classifier Systems; Belief function theory; Uncertainty; Classification; Machine Learning.

ACM Reference format:
Chedi Abdelkarim, Lilia Rejeb, Lamjed Ben Said, Maha Elarbi . 2017. Evidential Learning Classifier System. In *Proceedings of GECCO '17 Companion, Berlin, Germany, July 15-19, 2017,* 2 pages.
DOI: http://dx.doi.org/10.1145/3067695.3075997

1 INTRODUCTION

UCS proved to be efficient in classification. However, the classification in real world problems is characterized by uncertainty which could be present at data level (imprecision, incompleteness, etc) as well as class selection. In this paper, we are interested in uncertainty in class selection.

This kind of uncertainty was considered in two works that integrate fuzzy theory in accuracy-based algorithms: Fuzzy-UCS [4] and Fuzzy-XCS [2]. These studies require beforehand the transformation of real inputs rules into fuzzy rules. In this work, we propose to tackle uncertainty by avoiding the use of fuzzy rules. In fact, other theories could be used to deal with uncertainty such as probability theory, belief function theory and possibility theory. The belief function theory is a generalization of probability and possibility theories. It is a theory of quantified beliefs. It also provides an efficient framework to combine a variety of uncertain information [6].

Our method combines the sUpervised Classifier System (UCS) [1] and the Belief Function theory [5]. The new approach main goal is to improve the performance of UCS in the classification task. We aim to deal only with imprecision in the class selection.

2 THE BELIEF FUNCTION THEORY

The belief function theory was introduced by Dempster-Shafer [5]. It is also related to the fuzzy set theory [7], random sets [3], etc.

2.1 Frame of discernment

Θ is the frame of discernment. It is a finite non empty set that encompasses all the elementary events that appear in a particular problem. **Θ** is considered as the universe of discourse or the domain of reference.

In general, all the subsets of **Θ** belong to the power set of **Θ**, denoted by 2^{Θ} where an element of 2^{Θ} is referred as a proposition or an event.

2.2 Basic belief assignment

The basic belief assignment bba represents the belief attributed to the different subsets of the frame of discernment **Θ** which is defined as follows:

$$\begin{aligned} m : 2^{\Theta} &\rightarrow [0.1] \\ \textstyle\sum_{A\subseteq\Theta} m(A) &= 1 \end{aligned} \tag{1}$$

where **m(A)** is the basic belief mass **(bbm)**. It indicates the part of belief related to the event A of **Θ** given a piece of evidence. The focal element of a bba represents a strictly positive mass for every subset A of the frame of discernment **Θ**. **m(Θ)** quantifies the beliefs that are not attributed to any subsets of **Θ**.

2.3 Pignistic Probability

The TBM has to pass by a two level mental model: The first one is the credal level where beliefs are represented by belief function. The second one is the pignistic level which consists on transforming the beliefs into a probability in order to make decision. The latter called the pignistic probablity which is defined as follows:

$$BetP(A) = \sum_{B\subseteq\Theta} \frac{|A\cap B|}{|B|}\frac{m(B)}{1-m(\emptyset)}.\ for\ all \subseteq \Theta \tag{2}$$

3 AN EVIDENTIAL LEARNING CLASSIFIER SYSTEM

The Evidential Learning Classifier System starts by the initialization of its parameters.

GECCO '17 Companion, Berlin, Germany

DOI: http://dx.doi.org/10.1145/3067695.3075997

RUN ELCS (see Algorithm 1) begins when a new input is received. Run ELCS algorithm constructs the match set [M]. [M] contains the classifiers matching with i. It then applies the Exploration or the Exploitation strategy according to a certain threshold fixed a priori.

Algorithm 1 Run ELCS

```
[P] ← empty
repeat
    σ ←env:get situation
    [M] ← generate match set out of [P] considering the actual
    situation σ
     DoExploration(Envstate σ, MatchSet[M])
until the number of exploration steps is met
repeat
    σ ←env:get situation
    [M] ← generate match set out of [P] considering the actual
    situation σ
     DoExploitation(Envstate σ, MatchSet[M])
until the termination criterion is met
```

In exploration, the ELCS constructs the correct set [C] and the incorrect set [I]. If the match set is empty, then the covering is applied. Contrary to the exploration, the exploitation phase of ELCS consists in selecting the best class through the application of belief function theory. The Evidential Learning Classifier System reorganizes the formed match set [M] in groups of classifiers having the same proposed class. The set of these groups of classifiers is called [SetC]. The ELCS provides the masses of each class that belongs to the [SetC]. The masses are used to calculate the pignistic probability BetP (equation 2) which is used to select the action having the maximal value. If the ELCS is in the training phase, then, the parameters are updated and the GA is applied respecting to a fixed threshold.

4 EXPERIMENTATION

4.1 Experimental Protocol

We use data sets downloaded from UCI (the University of California at Irvine) repository in order to investigate the performance of the Evidential Learning Classifier System. We compare our approach also to those provided by the Weka toolbox such as : Support Vector Machine (SMO), Decision tree C4.5, PART, K-Nearest Neighbor (IB5). We adopt as a method of evaluation the 10-fold Cross-validation.

4.2 Results and Discussion

Table 1 shows the results of the classification of classical learning algorithms and ELCS. The last two rows represent respectively the average rank and the position for each learner. The experimental results show that the ELCS is one of the best ranked algorithms comparing to the other ones. The ELCS (2nd position) has an average rank very close to SMO (1st position). The ELCS has better performance than SMO for (bal, brt, gls, hcol, sht). For these data sets, the proposed approach improves significantly the accuracy by more than 2%. The proposed approach outperforms UCS where the difference between their position is 4. Also, it has higher performance comparing to C4.5 and PART which are rule-based algorithms. The ELCS is not challenging in the classification of few data sets (wdbc, wne) which is due to the high overlapping region between the classes. Also, these data sets have a large number of attributes (more than 10). So, a feature selection method should be adopted in such case.

[5] The choice of KNN parameter was based on the different experiments that was made with different k values {3,5,8}. Thus, we retain k=3 since it gives good results for the majority of data sets.

Table 1: The Performance of ELCS in term of classification accuracy comparing to classical machine learning algorithms.

ID	UCS	ELCS	C4.5	Part	SMO	IB3[5]
brt	65.32	68.0	65.09	64.15	59.43	66.03
cmc	50.27	52.38	53.02	50.23	51.39	46.03
gls	62.55	70.95	71.02	71.49	57.94	69.62
hts	74.63	82.31	74.44	75.18	84.44	79.62
hco	82.30	83.89	85.59	83.42	83.15	81.79
irs	95.32	95.72	95.33	95.33	95.33	96
pma	74.61	76.58	74.08	73.82	77.34	74.08
sht	99.62	99.50	99.97	99.97	96.96	99.90
vec	71.02	72.14	73.04	72.93	74.11	71.39
wbc	96.14	95.94	93.70	93.99	96.56	95.56
wdbc	95.52	94.63	93.49	93.14	97.89	96.66
wne	96.13	95.29	94.38	92.69	99.43	97.19
wpbc	69.4	77.37	76.66	76.26	77.27	73.73
Rank	3.37	2.56	2.87	3.06	2.25	2.75
Pos	6	***2***	4	5	***1***	3

CONCLUSION

We proposed in this paper, a new approach which takes account of uncertainty by integrating the Dempster-Shafer theory in the Learning Classifier System UCS. This new method accomplished a high performance in certain problems of classification comparing to UCS and to classical machine learning algorithms. However, it only takes account of uncertainty in the class selection. As future work, we will adapt the ELCS to deal with uncertainty in attributes and to treat the problem of incompleteness. We will also address the reduction of the number of rules to get more generalized patterns and to apply this new approach in the medical field.

REFERENCES

[1] Ester Bernadó-Mansilla and Josep M Garrell-Guiu. 2003. Accuracy-based learning classifier systems: models, analysis and applications to classification tasks. *Evolutionary computation* 11, 3 (2003), 209–238.
[2] Jorge Casillas, Brian Carse, and Larry Bull. 2007. Fuzzy-XCS: a michigan genetic fuzzy system. *IEEE Transactions on Fuzzy Systems* 15, 4 (2007), 536.
[3] Hung T Nguyen. 2006. *An introduction to random sets*. CRC press.
[4] Albert Orriols-Puig, Jorge Casillas, and Ester Bernadó-Mansilla. 2009. Fuzzy-UCS: a Michigan-style learning fuzzy-classifier system for supervised learning. *IEEE transactions on evolutionary computation* 13, 2 (2009), 260–283.
[5] Glenn Shafer and others. 1976. *A mathematical theory of evidence*. Vol. 1. Princeton university press Princeton.
[6] Philippe Xu. 2014. *Information fusion for scene understanding*. Ph.D. Dissertation. Compiègne.
[7] Lotfi A Zadeh. 1965. Fuzzy sets. *Information and control* 8, 3 (1965), 338–353.