

Apprentissage flou en environnement dynamique

Fuzzy Learning in Dynamical Environment

Christophe Marsala
UPMC Univ Paris 06, UMR 7606
LIP6, 4 place Jussieu, 75005, Paris
Christophe.Marsala@lip6.fr

Résumé :

Les environnements dynamiques sont de plus en plus présent dans notre société et offrent de nouveaux défis pour l'apprentissage automatique.

Nous nous plaçons ici dans un domaine où les données d'apprentissage sont fournies séquentiellement, voire temporellement, en un flux continu. Dans de tels environnements, les données dynamiques qui permettent de constituer les bases d'apprentissage peuvent être de plusieurs formes : données temporelles provenant de séries chronologiques, ou données résultant de mesures périodiques du phénomène étudié. Par exemple, on peut citer une série d'articles publiés dans des journaux en ligne (chaque document est alors décrit au moyen des thématiques ou des mots qui le constituent et qui évoluent au fil de l'actualité).

Toutefois, les évolutions dynamiques de ces données ne sont pas uniques et peuvent conduire à différents traitements. Par exemple, en apprentissage supervisé, la dynamité peut produire différents effets selon que le flux de données apporte une plus grande précision sur les frontières entre les classes, ou que les limites entre les classes évoluent au fil du temps : elles peuvent alors se déplacer dans l'espace de description, apparaître ou disparaître (ce peut être le cas lorsque la description des données évolue dans le temps comme dans l'exemple des journaux en ligne) [1, 3, 5].

Ainsi, de telles évolutions (on parle de "concept drift" ou de "concept shift" par exemple) gagnent à être prises en compte grâce à des approches d'apprentissage automatique flou [2] et, en particulier, par des approches par arbres de décision flous [4].

Mots-clés :

Apprentissage automatique flou. Systèmes adaptatifs.

Abstract:

Dynamical environments are more and more present in our society and offer a challenge for fuzzy machine learning. We place ourselves in a domain where the training data are provided sequentially, in a flow. In such environments, dynamic data that make up evolving training sets can be of several forms : temporal data coming from time series, or data resulting from periodical measurements of a studied phenomenon that evolves over time as, for instance, a set of papers published on online newspapers (each paper is described by means of the topics or words that evolve over time in the description space).

However, dynamic evolutions of such data are not unique

and can lead to different ways to handle them. For instance, in supervised learning, dynamicity occurs either when the flow of data provide more precision on the boundaries between the classes, or when the boundaries between the classes change over time: they could move in the description space, appear, or disappear (it could be the case when the description of the data evolve over time as in the online newspaper example) [1, 3, 5].

Thus, such concept drifts or shifts have to be handled and fuzzy machine learning [2] offers adequate tools in that task. In particular, fuzzy decision tree based learning is a promising approach for that task [4].

Keywords:

Fuzzy Machine Learning, Evolving Systems.

Références

- [1] P. Domingos and G. Hulten. Mining high-speed data streams. In *Proceedings of the KDD'00 Conference*, pages 71–80, 2000.
- [2] E. Hüllermeier. Fuzzy sets in machine learning and data mining. *Applied Soft Computing*, 11(2) :1493–1505, March 2011.
- [3] E. Lughofer and P. Angelov. Handling drifts and shifts in on-line data streams with evolving fuzzy systems. *Applied Soft Computing*, 11(2) :2057 – 2068, 2011.
- [4] C. Marsala. Fuzzy decision trees for dynamic data. In *IEEE Symposium on Evolving and Adaptive Intelligent Systems - EAIS'2013*, IEEE Symposium Series on Computational Intelligence, page to appear, Singapore, Apr. 2013.
- [5] G. Widmer and M. Kubat. Learning in the presence of concept drift and hidden contexts. *Machine Learning*, 23 :69–101, 1996.