Meta-Learning for Hierarchical Classification and Applications in Bioinformatics

Authors : Fabio Fabris, Alex A. Freitas

Abstract : Hierarchical classification is a special type of classification task where the class labels are organised into a hierarchy, with more generic class labels being ancestors of more specific ones. Meta-learning for classification-algorithm recommendation consists of recommending to the user a classification algorithm, from a pool of candidate algorithms, for a dataset, based on the past performance of the candidate algorithms in other datasets. Meta-learning is normally used in conventional, non-hierarchical classification. By contrast, this paper proposes a meta-learning approach for more challenging task of hierarchical classification, and evaluates it in a large number of bioinformatics datasets. Hierarchical classification is especially relevant for bioinformatics problems, as protein and gene functions tend to be organised into a hierarchy of class labels. This work proposes meta-learning approach for recommending the best hierarchical classification algorithm to a hierarchical classification dataset. This work’s contributions are: 1) proposing an algorithm for splitting hierarchical datasets into new datasets to increase the number of meta-instances, 2) proposing meta-features for hierarchical classification, and 3) interpreting decision-tree meta-models for hierarchical classification algorithm recommendation.
Keywords : algorithm recommendation, meta-learning, bioinformatics, hierarchical classification

Conference Title : ICDM 2018 : International Conference on Data Mining
Conference Location : Dublin, Ireland
Conference Dates : July 23-24, 2018