Active Learning Using Uncertainty Information

Yang, Yazhou, and Marco Loog. "Active learning using uncertainty information." *2016 23rd International Conference on Pattern Recognition (ICPR)*. **IEEE, 2016.**

What problem does this paper try to solve?

This paper proposes a different method to enhance the performance of existing retraining-based models. Existing methods resort to calculating the average- or worst-case performance regarding unknown labels. The paper also incorporates the uncertainty information and applies the proposed method to two state-of-the-art algorithms such as EER(Expected Error Reduction) and MLI(Minimum Loss Increase), showing the effectiveness of the proposed method.

How does it solve the problem?

Not knowing the true labels of the unlabeled data may result in failure to take account some potentially valuable information. The paper introduces a new criterion for retraining-based active learning to overcome the limitations of existing methods.

The new criterion also incorporates uncertainty information within the min-max framework, balancing exploration and exploitation. The method can be applied to two state-of-the-art algorithms: Expected Error Reduction (EER) and Minimum Loss Increase (MLI).

List of novelties/contributions

- It proposes a new active learning criterion that incorporates uncertainty information with retraining-based models.

- It extends existing methods, EER and MLI, by applying the proposed criterion, creating uncertainty-based versions, UEER and UMLI.

- It shows that the new method balances exploration and exploitation by incorporating uncertainty information and improves the performance of retraining-based models.

- It presents extensive experiments on 49 real-world datasets showing that the new method outperforms existing approaches in most cases.

Downsides of the work

- The proposed method still requires retraining on multiple examples, which can be computationally expensive in large datasets.

- From the results of the experiments, performance of the proposed method may not be as good as the performance of existing methods in some cases.