We introduce Information Forests, an approach to classification that generalizes Random Forests by replacing the splitting criterion of non-leaf nodes from a discriminative one – based on the entropy of the label distribution – to a generative one – based on maximizing the information divergence between the class-conditional distributions in the resulting partitions. The basic idea consists of deferring classification until a measure of “classification confidence” is sufficiently high, and instead breaking down the data so as to maximize this measure. In an alternative interpretation, Information Forests attempt to partition the data into subsets that are “as informative as possible” for the purpose of the task, which is to classify the data. Classification confidence, or informative content of the subsets, is quantified by the Information Divergence. Our approach relates to active learning, semi-supervised learning, mixed generative/discriminative learning.

Topic: active learning
Preference: poster