**Active Learning by Labeling Features**

Year: 2009

<https://pdfs.semanticscholar.org/54d2/be3b053c36b7b8fb928926c19da609143be2.pdf>

Propose an active learning approach in which the machine solicits “labels” on features rather than instances.

*Methods:*

Label instance:

ER: entropy regularization

Label features:

Using generalized expectation (GE) criteria for learning with labeled features.

A pool-based feature active learning algorithm that includes an option to skip queries, for cases in which a feature has no clear label.

Propose and evaluate feature query selection algorithms that aim to reduce model uncertainty and compare to several baselines.

*Feature query selection methods φ:*

* Expected Information Gain: the expectation of the reduction in model uncertainty over all possible responses.
* Total Uncertainty: the sum of the marginal entropies at the positions where the feature occurs.
* Weighted Uncertainty: scales the mean uncertainty by the log count of the feature in the corpus.
* Diverse Uncertainty: encourages diversity among queries by multiplying TU by the mean dissimilarity between the feature and previously labeled features.
* Coverage: aims to select features that are dissimilar from existing labeled features, increasing the labeled features’ “coverage” of the feature space.
* Similarity: the maximum similarity to a labeled feature, weighted by the log count of the feature.
* Random: assigns scores to features randomly.
* Frequency: scores input features using their frequency in the training data.

When presented an instance query, the oracle simply provides the true labels. When presented a feature query, the oracle first decides whether to skip the query.

**Fine-tuning Convolutional Neural Networks for Biomedical Image Analysis: Actively and Incrementally**

Year: 2017

<http://openaccess.thecvf.com/content_cvpr_2017/papers/Zhou_Fine-Tuning_Convolutional_Neural_CVPR_2017_paper.pdf>

Git:

<https://github.com/MrGiovanni/Active-Learning>

*How to cut annotation cost?*

Motivation: Annotating biomedical images is very challenging. It is not only tedious and time

consuming, but also demanding of costly, specialty-oriented knowledge and skills, which are not easily accessible.

*The method (improved active learning)*:

AIFT: active incremental fine-tuning

*Key Ideas:*

Active Learning + Transfer Learning

Data Augmentation

Majority Selection

Continuously Fine-Tuning

Entropy + Diversity

1. Active selection: consistency among the patches generated from a candidate.

2. Handling noisy labels: majority selection.

3. Continuous fine-tuning: fine-tuning the fine-tuned CNN.

Evaluated the method in three different biomedical imaging applications, including colonoscopy frame classification, polyp detection, and pulmonary embolism (PE) detection, demonstrating that the cost of annotation can be cut by at least half.

*Advantages:*

1. Starting with a completely empty labeled dataset.

2. Incrementally improving the learner through continuous fine-tuning rather than repeatedly re-training.

3. Naturally exploiting expected consistency among the patches associated for each candidate to select samples “worthy” of labeling.

4. Automatically handling noisy labels as only a portion of the patches in each candidate participates in the selection process.

5. Computing entropy and diversity locally on a small number of patches within each candidate, saving computation time considerably.

**Active learning by querying informative and representative examples**

Year: 2010

<https://papers.nips.cc/paper/4176-active-learning-by-querying-informative-and-representative-examples.pdf>

Most active learning approaches select either informative or representative unlabeled instances to query their labels.

*The method:*

A new active learning approach by Querying Informative and Representative Examples (QUIRE for short).

The proposed approach is based on the min-max view of active learning, which provides a systematic way for measuring and combining the informativeness and the representativeness of an instance.

Informativeness measures the ability of an instance in reducing the uncertainty of a statistical model, while representativeness measures if an instance well represents the overall input patterns of unlabeled data.

*Experiments:*

Compare QUIRE with the following five baseline approaches:

(1) RANDOM: randomly select query instances

(2) MARGIN: margin-based active learning, a representative approach which selects informative instances

(3) CLUSTER: hierarchical-clustering-based active learning, a representative approach that chooses representative instances

(4) IDE: active learning that selects informative and diverse examples

(5) DUAL: a dual strategy for active learning that exploits both informativeness and representativeness for query selection.

**Multi-Class Active Learning by Uncertainty Sampling with Diversity Maximization**

Year: 2014

<https://link.springer.com/content/pdf/10.1007%2Fs11263-014-0781-x.pdf>

Methods:

Uncertainty sampling with diversity maximization

Most existing active learning algorithms only exploit the labelled data, which often suffers from over-fitting due to the small number of labelled examples.

A semi-supervised batch mode **multi-class** active learning algorithm for visual concept recognition. This algorithm exploits the whole active pool to evaluate the uncertainty of the data. Propose to make the selected data as diverse as possible, for which we explicitly impose a diversity constraint on the objective function.

As a multi-class active learning algorithm, this algorithm is able to exploit uncertainty across multiple classes. An efficient algorithm is used to optimize the objective function. Extensive experiments on action recognition, object classification, scene recognition, and event detection demonstrate its advantages.

**Learning Active Learning from Data**

Year: 2017

<https://papers.nips.cc/paper/7010-learning-active-learning-from-data.pdf>

Github:

<https://github.com/ksenia-konyushkova/LAL>

Suggests a data-driven approach to active learning that trains a secondary model to identify the unlabeled data points which, when labeled, would likely have the greatest impact on our primary model’s performance.