CSE 584, Fall 2024

Homework 1

Name: Samira Malek **PSU ID: sxm6547**

Task description of HW1: Find and read three papers about Active Learning. For each paper, please write down:

1. What problem does this paper try to solve, i.e., its motivation
2. How does it solve the problem?
3. A list of novelties/contributions
4. What do you think are the downsides of the work?

First paper: How to Select Which Active Learning Strategy is Best Suited for Your Specific Problem and Budget

---

Summary:

The paper addresses the problem of selecting the most appropriate active learning strategy for a specific problem and budget, a challenge that researchers often face when working with machine learning models. In active learning, a model selectively queries data points from an oracle to be labeled, aiming to maximize performance while minimizing labeling costs. However, choosing the right AL strategy can be difficult because the effectiveness of a strategy depends heavily on the budget and the nature of the problem.

1. What problem does this paper try to solve, i.e., its motivation

The paper tackles the challenge of determining the best Active Learning strategy for a specific task and budget constraints. The motivation behind the study is the observation that different AL strategies work better under different budget conditions—uncertainty-based methods perform well with larger budgets, while typicality-based methods are more suitable for smaller budgets.

The challenge lies in the fact that in real-world scenarios, determining in advance which AL strategy will be most effective for a given problem and budget is difficult. The traditional approaches do not dynamically adapt to varying budget sizes and problem characteristics, often leading to suboptimal performance. This paper seeks to solve this problem by introducing a method that can automatically select the best AL strategy based on the characteristics of the problem and the budget, without requiring extensive prior knowledge.

2. How does it solve the problem?

The paper proposes **SelectAL**, a practical algorithm to solve the problem of selecting the most suitable Active Learning strategy for a given budget and problem setting.

The authors first develop a simplified theoretical framework to analyze how different AL strategies perform depending on the budget. They demonstrate that low-budget strategies (focused on typical examples) are more effective when the labeling budget is small, while high-budget strategies (focused on uncertain or diverse examples) work better with larger budgets.

SelectAL has two parts:
1: The algorithm evaluates the current labeled set by applying small data perturbations using both low-budget and high-budget AL strategies. It compares how the model's error changes when using each strategy to determine the budget regime.
2: Based on the identified budget regime, SelectAL chooses the appropriate AL strategy for the next round of labeling. If the budget is small, it selects a typicality-based strategy; if the budget is large, it selects an uncertainty-based strategy.

3. A list of novelties/contributions
1: Their method can adapt AL strategies to different budget constraints, providing a versatile solution for a wide range of budgets.
2: A key contribution is the derivative-based test, which uses small data perturbations to estimate whether the current budget should be treated as "small" or "large."
3: SelectAL can dynamically shift between low-budget and high-budget AL strategies during different stages of the labeling process, ensuring optimal performance across varying budget sizes.

4. What do you think are the downsides of the work?
The theoretical framework used to develop the derivative-based test is highly simplified. While it provides clear insights and motivations, it may not fully capture the complexities of real-world AL scenarios, especially in deep learning tasks. The assumptions made in the framework might not generalize well to more complex and diverse datasets or tasks beyond those tested in the paper. SelectAL relies heavily on its ability to classify the budget as either “small” or “large.” There may be cases where this classification is ambiguous, especially when budgets fall into mid-range categories. This could lead to suboptimal strategy selection if the algorithm misclassifies the budget.

Second paper: (TiDAL) Learning Training Dynamics for Active Learning

----

1. What problem does this paper try to solve, i.e., its motivation

This paper tries to solve the inefficiency of uncertainty estimation in active learning methods. Most traditional AL methods, especially uncertainty-based methods, rely on static snapshots of the model after training to estimate uncertainty, such as using the predicted probability distributions or loss values at the end of training. However, this approach misses valuable information that is generated throughout the training process (training dynamics (TD)).

The motivation of the paper is based on the idea that training dynamics — the changing behavior of the model over time during optimization via stochastic gradient descent — provide richer information about the uncertainty of data samples. Recent research has shown that tracking how a model learns a sample over time can offer important clues about whether that sample is certain (easy) or uncertain (hard) for the model. However, tracking TD for all data samples during training is computationally prohibitive, especially when dealing with large unlabeled data pools in AL.

2. How does it solve the problem?

This paper introduced a novel method that leverages training dynamics (TD) to estimate the uncertainty of unlabeled data samples in a computationally efficient way. The following is a summary of their method:

First, a target classifier is trained on a small set of labeled data, and the TD of these labeled samples is monitored throughout the training process. Next, a TD prediction module is developed using the TD data from the labeled samples, enabling the module to learn how to predict TD for new, unseen samples. This allows the TD prediction module to estimate the TD of unlabeled samples in the pool without needing to infer their TD at every training epoch. The uncertainty of each unlabeled sample is then estimated using the predicted TD along with standard uncertainty measures, such as entropy or margin. Finally, the most uncertain samples are selected based on these estimates and sent to human annotators for labeling.

3. A list of novelties/contributions

1: linking TD with active learning (AL), proposing that TD can better estimate the uncertainty of unlabeled data compared to traditional static methods that rely on a fully trained model's snapshot.

2: To overcome the high computational cost of tracking TD for all unlabeled data, the paper introduces a TD prediction module. This module predicts the TD of unlabeled data based on the TD of labeled data, which is readily available during training, making the process computationally efficient.

4. What do you think are the downsides of the work?

1: TiDAL relies on the outputs of the target classifier to predict training dynamics (TD). If the classifier performs poorly, especially in early training stages, it may misclassify hard samples with high confidence, treating them as certain. This could lead to suboptimal sample selection in the active learning process, as the model might overlook important, uncertain samples that need labeling.

2: TiDAL is specifically designed for multi-class classification tasks and may not be applicable to other types of tasks, such as regression or more complex domains like object detection or segmentation. This restricts its broader applicability to tasks outside of classification.

Third paper: Hybrid Active Learning via Deep Clustering for Video Action Detection

---

1. What problem does this paper try to solve, i.e., its motivation

The paper is motivated by the high annotation costs associated with video action detection tasks, which require detailed, frame-wise spatio-temporal annotations. Unlike video classification tasks, where only video-level annotations are sufficient, video action detection demands annotations for every frame, including bounding boxes or pixel-wise labels, significantly increasing the labeling burden and cost.

2. How does it solve the problem?

The paper solves the problem of high annotation costs in video action detection through a hybrid active learning strategy that combines both intra-sample (frame-level) and inter-sample (video-level) selection. This approach leverages Clustering-Aware Uncertainty Scoring (CLAUS) to ensure informative and diverse sample selection, while the Spatio-Temporal Weighted (STeW) loss enhances model training under limited annotations by utilizing temporal continuity and pseudo-labels. By selecting only the most important frames and videos, the method significantly reduces annotation redundancy and cost, while maintaining performance comparable to fully-supervised models. Through iterative active learning cycles, the model gradually improves with minimal annotations, effectively balancing performance and cost efficiency.

3. A list of novelties/contributions

1: A novel combination of intra-sample (frame-level) and inter-sample (video-level) selection, reducing the overall annotation cost by selecting informative and diverse frames and videos.
2: A new sample selection approach that integrates model uncertainty (informativeness) with deep clustering (diversity), ensuring that selected samples are both informative and non-redundant.

4. What do you think are the downsides of the work?

1: The hybrid active learning framework, combining intra-sample and inter-sample selection with clustering and uncertainty scoring, adds a layer of complexity to the overall system. Implementing and fine-tuning such a system may be more challenging compared to simpler active learning methods.
2: The effectiveness of the Clustering-Aware Uncertainty Scoring (CLAUS) heavily depends on the model's uncertainty estimates, which can sometimes be unreliable, particularly in early training stages when the model is less accurate. This may lead to suboptimal sample selections.