CSE-584 Homework #1

Literature Review of Active Learning

Paper #1

"A Simple Baseline for Low-Budget Active Learning"

1. Problem Statement

   This paper highlights that most active learning methods assume access to a large subset of annotated data. In contrast, the authors propose a low-budget active learning approach where only 0.2% of ImageNet is annotated. They demonstrate that by using a simple K-means clustering algorithm with limited resources, it is possible to outperform state-of-the-art active learning methods that rely on larger labeling budgets.

2. How does it solve the problem (Method)

   They utilized a simple K-means clustering algorithm and their approach leverages rich feature extraction using an off-the-shelf self-supervised learning method applied once, followed by an evaluation of different sampling strategies on various datasets, including ImageNet, under a low-labeling budget scenario.

3. Novelties/Contributions
   a. The team utilized a simple K-means clustering algorithm to outperform the state-of-the-art active learning methods with low budget.
   b. They explore two forms of K-means sampling for active learning in low-budget scenarios:
      i. Single-batch K-means: This method performs only one iteration of sampling, applying K-means clustering to feature outputs from a self-supervised learning (SSL) pre-trained model. Samples closest to the cluster centers are selected for annotation. It simplifies the annotation workflow as annotators do not need to wait for additional model training between batches. However, it requires knowledge of the total number of samples that need annotation upfront.
      ii. Multi-batch K-means: This iterative method selects new samples for annotation across multiple rounds. It computes clusters based on the difference between consecutive budget sizes and selects unlabeled examples closest to the new cluster centers. Although it avoids wasting previously labeled data, it requires waiting for prior rounds to complete, making it more time-intensive.
4. Potential Downsides of the Work

   The team explored two methods, single-batch and multi-batch K-means, but the results showing single-batch has better performance than the multi-batch which was not what they expected. The paper assumes this happens because single-batch method finds a large number of clusters that matches the total budget, which causes to represent very small clusters as well instead of performing experiments to prove this. However, this work still proposed a strong active learning baseline to achieve an accurate image classifier in very low budgets as they mentioned.

Paper #2

"Towards Robust and Reproducible Active Learning using Neural Networks"

1. Problem Statement

   Recently proposed neural network-based active learning (AL) methods use various heuristics to achieve their goals. In this study, the team demonstrate that, under identical experimental conditions, different types of AL algorithms (uncertainty-based, diversity-based, and committee-based) show inconsistent improvements over the random sampling baseline. Through extensive experiments controlling for sources of stochasticity, they revealed that the variance in performance metrics of AL algorithms can lead to results that contradict previously reported findings. Additionally, the team found that under strong regularization, AL methods offer only marginal or no advantage over random sampling across various experimental conditions.

2. How does it solve the problem (Method)

   The team conclude by providing a set of recommendations for assessing the results of new active learning (AL) algorithms to ensure their reproducibility and robustness under varying experimental conditions. These guidelines focus on controlling for stochasticity, comparing against strong baselines, and evaluating performance across a diverse range of datasets and scenarios to ensure consistent and reliable outcomes

3. Novelties/Contributions
   a. Pool Based Active Learning Methods
   b. Model Uncertainty on Output
   c. Deep Bayesian Active Learning
   d. Variational Adversarial Active Learning
   e. Ensemble Variance Ratio Learning
   f. The paper provides very sufficient experiments and discussion about the topic
4. Potential Downsides of the Work
   Although the team provides very carefully considered experiments, they only conducted for image classification without further explore on detection and segmentation as they also mentioned in the paper.

Paper #3

"Deep Bayesian Active Learning for Preference Modeling in Large Language Models"

1. Problem Statement
   Leveraging human preferences to guide the behavior of Large Language Models (LLMs) has shown significant success in recent years. However, data selection and labeling remain bottlenecks, especially at large scales. Therefore, selecting the most informative

data points for human feedback can significantly reduce the cost of preference labeling and accelerate the development of LLMs. Bayesian Active Learning offers a principled framework to address this challenge and has proven successful in various applications. However, prior efforts to apply it to Preference Modeling have fallen short of these expectations.

2. How does it solve the problem (Method)
   In this work, they observed that naive epistemic uncertainty estimation often leads to the acquisition of redundant samples. To address this, the team introduce the Bayesian Active Learner for Preference Modeling (BAL-PM), a novel stochastic acquisition policy that not only targets points with high epistemic uncertainty according to the preference model but also aims to maximize the entropy of the acquired prompt distribution within the feature space of the LLM. Our experiments show that BAL-PM reduces the need for preference labels by 33% to 68% in two popular human preference datasets and outperforms previous stochastic Bayesian acquisition policies.
3. Novelties/Contributions
   a. Preference Modeling: Their model follows the Bradley-Terry(BT) model, the BT model is often implemented by learning a parameterized latent reward model and optimizing theta via maximum likelihood estimation. Which minimizing the negative log-likelihood with respect to the human preference labels.
   b. The team also adopt a Bayesian Model, which assumes a probability distribution over the parameters theta, they followed the Bayesian Active Learning by Disagreement (BALD) for active learning.
   c. Kozachenko-Leonenko Entropy
   d. The team proposed BAL-PM, a Bayesian Active Learning method for Preference Modeling in Language Models. BAL-PM substantially reduces the volume of feedback required for Preference Modeling and outperforms existing Bayesian stochastic acquisition policies.
4. Potential Downsides of the Work

   As the paper also mentioned, BAL-PM is heavily relying on the quality of the feature representations that the base LLM offered, however, this limitation seems to be addressed by progressing of LLMs these days.

Reference

Pourahmadi, Kossar, Parsa Nooralinejad, and Hamed Pirsiavash. "A simple baseline for low-budget active learning." arXiv preprint arXiv:2110.12033 (2021).

Munjal, Prateek, et al. "Towards robust and reproducible active learning using neural networks." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.

Melo, Luckeciano C., et al. "Deep Bayesian Active Learning for Preference Modeling in Large Language Models." *arXiv preprint arXiv:2406.10023* (2024).

Pourahmadi, Kossar, Parsa Nooralinejad, and Hamed Pirsiavash. "A simple baseline for low-budget active learning." arXiv preprint arXiv:2110.12033 (2021).