

HW1 - Analyzing three paper about active learning

Name: Janice Ahn

Date: 9/8/2024

1. Active Learning From Imbalanced Data: A Solution of Online Weighted Extreme Learning Machine (doi: 10.1109/TNNLS.2018.2855446)

- a. **Authors:** Hualong Yu; Xibei Yang; Shang Zheng; Changyin Sun
- b. **What problem:** The effectiveness of active learning is often hindered by imbalanced data distributions. Some current imbalanced active learning methods either perform poorly or are too time-consuming.
- c. **How does it solve the problem?**
 - i. By providing experiments with their claiming method
- d. **Contributions:** Introducing active online-weighted ELM (AOW-ELM) as solution;
 - i. Discusses how imbalanced instance distribution disrupts active learning.
 - ii. Uses hierarchical clustering to improve initial instance labeling.
 - iii. Proposes weighted ELM (WELM) as a fair base classifier with an online update mode.
 - iv. Introduces a flexible early stopping criterion.
- e. **Downside:** The AOW-ELM algorithm is particularly vulnerable to noise, especially in datasets with a high level of imbalance, where random sampling (ROW-ELM) can occasionally achieve better results than uncertainty-based sampling.

2. Supervised machine learning and active learning in classification of radiology reports

(<https://doi.org/10.1136/amiajnl-2013-002516>)

- a. **Authors:** Dung H M Nguyen, Jon D Patrick
- b. **What problem:** They need a survey of optimal data selection techniques
- c. **How does it solve the problem?**
 - i. They collected previous researches and analyzed them in terms of optimal data selection techniques
- d. **Contributions:**
 - i. Explained Active Learning as a Statistical Approach with examples
 - ii. Analyzing Active Learning with a Mixtures of Gaussians which is a powerful estimation and prediction technique model.
 - iii. Locally Weighted Regression for Active Learning was provided for analysis
- e. **Downside:** They focused solely on function approximation problems.

3. AL-ELM: One uncertainty-based active learning algorithm using extreme learning machine (<https://doi.org/10.1016/j.neucom.2015.04.019>)

- a. **Authors:** Hualong Yu, Changyin Sun, Wankou Yang, Xibei Yang, Xin Zuo
- b. **What problem:** labeling instances is not cost / time efficient
- c. **How does it solve the problem?**
 - i. Using extreme learning machine (ELM) classifier
- d. **Contributions:**
 - i. Introducing ELM which has low computational costs and strong generalization ability. Also, ELM can also be applied to either binary-class and multiclass problems.
 - ii. They conducted 20 benchmark datasets experiment with active learning algorithm based on ELM (AL-ELM), which showed a result of reducing running time of learning procedure
- e. **Downside:** It would be better to have more comparison with other methods not only with the non-uncertainty-based active learning techniques or at least mentioning it would be great