A Comparative Survey of Deep Active Learning

arXiv 2022

Xueying Zhan∗ Qingzhong Wang Kuan-Hao Huang
Haoyi Xiong Dejing Dou Antoni B. Chan

- Problem / objective
  - Several query strategies in active learning
- Contribution / Key idea
  - Several query strategies in active learning

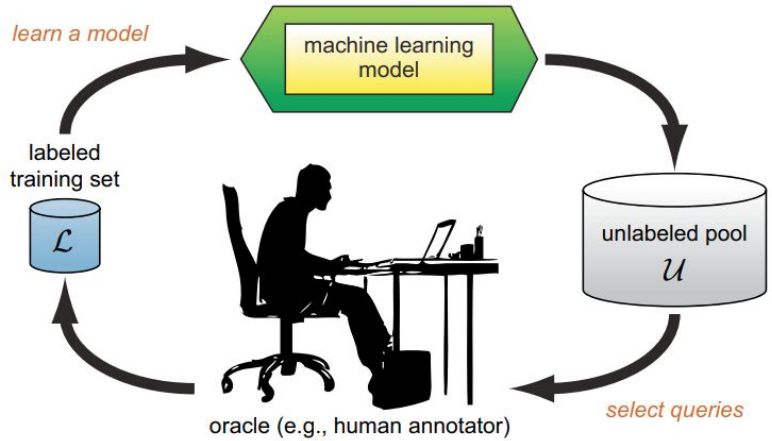
A Comparative Survey of Deep Active Learning

*Xueying Zhan** *Qingzhong Wang* *Kuan-Hao Huang*
Haoyi Xiong *Dejing Dou* *Antoni B. Chan*

arXiv 2022

Active Learning

- labeling cost 비싼 문제 해결하기 위해 등장한 학습 방법.
- 도움이 될것같은 최소한의 데이터만 라벨링하여 학습.



A Comparative Survey of Deep Active Learning

arXiv 2022

*Xueying Zhan** *Qingzhong Wang* *Kuan-Hao Huang*
Haoyi Xiong *Dejing Dou* *Antoni B. Chan*

Overview

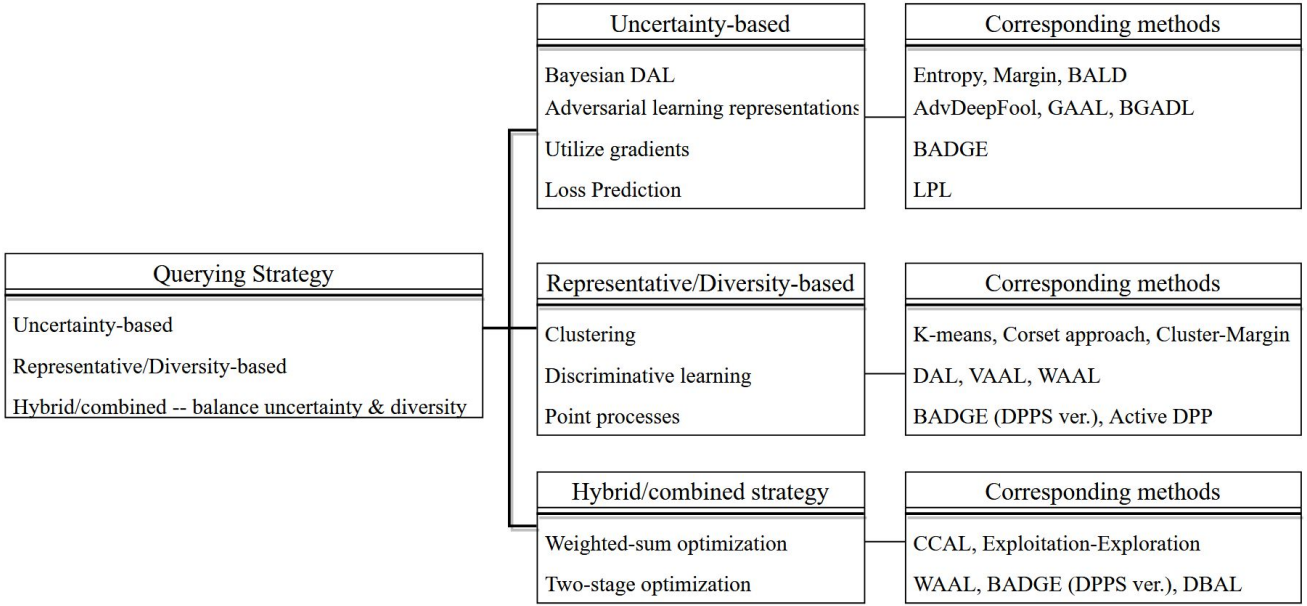


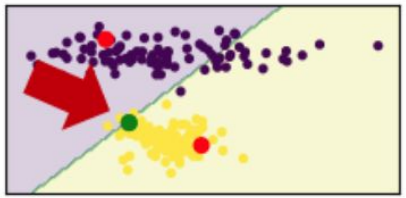
Figure 1: Categorization of DAL sampling/querying strategies.

A Comparative Survey of Deep Active Learning

*Xueying Zhan** *Qingzhong Wang* *Kuan-Hao Huang*
Haoyi Xiong *Dejing Dou* *Antoni B. Chan*

arXiv 2022

Uncertainty-Based Sampling Strategies



- 모델이 헷갈려하는 데이터 샘플링

1. Maximum Entropy
   a. 엔트로피가 큰 데이터 선택

$$\alpha_{\textbf{entropy}}(\mathbf{x}, \mathcal{M}) = H_{\mathcal{M}}[y|\mathbf{x}] = -\sum_k p_{\mathcal{M}}(y = k|\mathbf{x}) \log p_{\mathcal{M}}(y = k|\mathbf{x})$$

2. Margin
   a. 마진이 작은 데이터 선택

$$\alpha_{\textbf{margin}}(\mathbf{x}, \mathcal{M}) = -[p_{\mathcal{M}}(\hat{y}_1|\mathbf{x}) - p_{\mathcal{M}}(\hat{y}_2|\mathbf{x})]$$

3. Least Confidence
   a. Top-1 confidence 가 가장 낮은 데이터 선택

$$\alpha_{\textbf{LeastConf}}(\mathbf{x}, \mathcal{M}) = -p_{\mathcal{M}}(\hat{y}|\mathbf{x})$$

   b. Variation Ratios

$$\alpha_{\textbf{VarRatio}}(\mathbf{x}, \mathcal{M}) = 1 - p_{\mathcal{M}}(\hat{y}|\mathbf{x})$$

전유진

A Comparative Survey of Deep Active Learning

*Xueying Zhan** *Qingzhong Wang* *Kuan-Hao Huang*
Haoyi Xiong *Dejing Dou* *Antoni B. Chan*

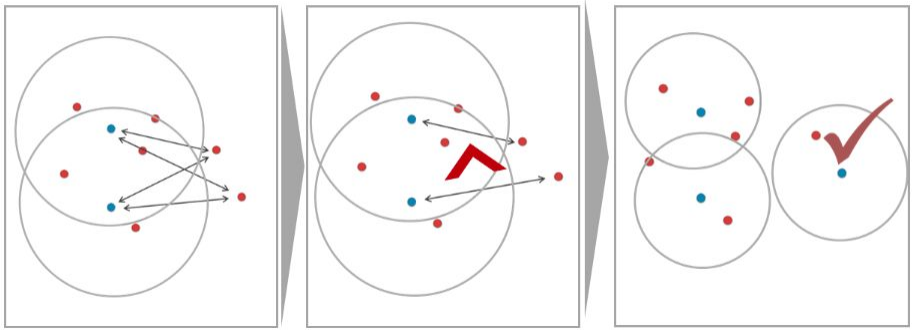
arXiv 2022

Diversity-Based Sampling Strategies

- 비슷한 샘플들 최대한 피해 다양한 데이터 샘플링

1. KMeans
   a. 각 클러스터의 중심점 선택

2. CoreSet https://arxiv.org/abs/1708.00489
   a. 기존 레이블 데이터와의 최소거리가 최대인 데이터 선택
   b. 대표적으로, K-Center-Greedy algorithm



3. Cluster-Margin https://arxiv.org/abs/2107.14263
   a. 각 클러스터 내에서 마진이 작은 데이터 선택

A Comparative Survey

*Xueying Zhan** *Qingz*
Haoyi Xiong *Dejing*

arXiv 2022

Experiments

	Model	*MNIST* AUBC	*MNIST* F-acc	*FashionMNIST* AUBC	*FashionMNIST* F-acc	*EMNIST* AUBC	*EMNIST* F-acc	*SVHN* AUBC	*SVHN* F-acc	*PneumoniaMNIST* AUBC	*PneumoniaMNIST* F-acc
	Full	–	0.9916	–	0.9120	–	0.8684	–	0.9190	–	0.9039
	Random	0.9570	0.9738	0.8313	0.8434	0.8057	0.8377	0.8110	0.8806	0.8283	**0.9077**
Unc	**LeastConf**	0.9677	0.9892	0.8377	0.8820	0.8113	0.8479	0.8350	0.9094	0.8520	**0.9097**
Unc	**LeastConfD**	0.9750	0.9915	0.8450	0.8744	0.8117	0.8483	0.8320	0.9083	0.8243	0.8654
Unc	**Margin**	0.9733	0.9881	0.8427	0.8772	0.8103	0.8468	0.8373	0.9138	0.8580	0.8859
Unc	**MarginD**	0.9703	0.9899	0.8417	0.8756	0.8197	0.8472	0.8357	0.9104	0.8230	**0.9149**
Unc	**Entropy**	0.9723	0.9883	0.8397	0.8660	0.8090	0.8458	0.8297	0.9099	0.8570	**0.9132**
Unc	**EntropyD**	0.9683	0.9887	0.8417	0.8784	0.8167	0.8507	0.8290	0.9091	0.8177	0.8710
Unc	**BALD**	0.9697	0.9885	0.8423	0.8888	0.8197	0.8448	0.8333	0.9020	0.8270	**0.9204**
Unc	**MeanSTD**	0.9713	0.9735	0.8457	0.8766	0.8110	0.8426	0.8323	0.9087	0.7827	0.8802
Unc	**VarRatio**	0.9717	0.9841	0.8410	0.8754	0.8107	0.8497	0.8357	0.9079	0.8530	0.8672
Unc	**CEAL(Entropy)**	0.9787	0.9889	0.8477	0.8826	0.8163	0.8459	0.8430	0.9142	0.8543	**0.9179**
Repr/Div	**KMeans**	0.9640	0.9813	0.8260	0.8525	0.7903	0.8264	0.8027	0.8671	0.8243	**0.9044**
Repr/Div	**KMeans (GPU)**	0.9637	0.9747	0.8343	0.8657	0.7990	0.8362	0.8120	0.8688	0.8333	**0.9155**
Repr/Div	**KCenter**	0.9740	0.9877	0.8353	0.8466	*	*	0.8283	0.9000	0.8130	**0.9189**
Repr/Div	**VAAL**	0.9623	0.9573	0.8297	0.8535	0.8027	0.8363	0.8117	0.8813	0.8393	0.9064
	BADGE(KMeans++)	0.9707	0.9904	0.8437	0.8662	*	*	0.8377	0.9057	0.8340	**0.9066**
Enhance	**AdvBIM**	0.9680	0.9840	0.8437	0.8729	#	#	#	#	0.8297	0.9197
Enhance	**LPL**	0.8913	0.9732	0.7600	0.8471	0.5640	0.6474	0.8737	**0.9452**	0.8593	**0.9346**
Enhance	**WAAL**	0.9890	**0.9946**	0.8703	0.8984	0.8293	0.8423	0.8603	0.9135	0.9663	**0.9564**
		CIFAR10		*CIFAR100*		*CIFAR10-imb*		*Tiny ImageNet*		*BreakHis*	
	Full	–	0.8793	–	0.6062	–	0.8036	–	0.4583	–	0.8306
	Random	0.7967	0.8679	0.4667	0.5903	0.7103	**0.8105**	0.2577	0.3544	0.8010	0.8150
Unc	**LeastConf**	0.8150	0.8785	0.4747	0.6072	0.7330	0.8022	0.2417	0.3470	0.8213	0.8302
Unc	**LeastConfD**	0.8137	**0.8825**	0.4730	0.5997	0.7323	**0.8065**	0.2620	0.3698	0.8140	**0.8313**
Unc	**Margin**	0.8153	**0.8834**	0.4790	0.6010	0.7367	0.8029	0.2557	0.3611	0.8217	0.8289
Unc	**MarginD**	0.8140	0.8837	0.4777	0.6000	0.7260	**0.8128**	0.2607	0.3541	0.8253	**0.8364**
Unc	**Entropy**	0.8130	0.8784	0.4693	0.6048	0.7320	0.8187	0.2343	0.3346	0.8213	0.8251
Unc	**EntropyD**	0.8140	0.8787	0.4677	0.6004	0.7317	0.7963	0.2627	0.3716	0.8017	0.8115
Unc	**BALD**	0.8103	0.8762	0.4760	0.5942	0.7210	0.7927	0.2623	0.3648	0.8147	0.8296
Unc	**MeanSTD**	0.8087	**0.8821**	0.4717	0.5963	0.7203	0.7996	0.2510	0.3551	0.8053	0.8202
Unc	**VarRatio**	0.8150	08780	0.4747	0.5959	0.7353	**0.8165**	0.2407	0.3426	0.8197	0.8264
Unc	**CEAL(Entropy)**	0.8150	**0.8794**	0.4693	0.6043	0.7327	0.8187	0.2347	0.3400	0.8163	0.8181
Repr/Div	**KMeans**	0.7910	0.8713	0.4570	0.5834	0.7070	0.7908	0.2447	0.3385	0.8203	0.8394
Repr/Div	**KMeans (GPU)**	0.7977	0.8718	0.4687	0.5842	0.7140	0.7921	0.1340	0.2288	0.8140	**0.8323**
Repr/Div	**KCenter**	0.8047	0.8741	0.4770	0.5993	0.7233	0.7826	0.2540	0.346	0.8027	0.8289
Repr/Div	**VAAL**	0.7973	0.8679	0.4693	0.5870	0.7113	0.7950	0.1313	0.2191	0.8197	**0.8344**
	BADGE(KMeans++)	0.8143	**0.8794**	0.4803	0.6034	0.7347	**0.8126**	#	#	0.8343	0.8470
Enhance	**AdvBIM**	0.7997	0.8750	0.4713	0.5855	#	#	#	#	0.8240	**0.8337**
Enhance	**LPL**	0.8220	**0.9028**	0.4640	**0.6369**	0.7477	**0.8478**	0.0090	0.0051	0.8277	**0.8316**
Enhance	**WAAL**	0.8253	0.8717	0.4277	0.5560	0.7523	0.7993	0.0157	0.0050	0.8620	**0.8698**

Table 1: Overall results of DAL comparative experiments. We **bold** F-acc values that are higher than full performance. We rank F-acc and AUBC of each task with top 1st, 2nd and 3rd with red, teal and blue respectively. "*" refers to the experiment needed too much memory, *e.g.*, **KCenter** on *EMNIST*. "#" refers to the experiment that has not been completed yet. Completed tables of all tasks are shown in Tables 4, 5, 6, 7, and 8 in Appendix.

A Comparative Survey of Deep Active Learning

*Xueying Zhan** *Qingzhong Wang* *Kuan-Hao Huang*
Haoyi Xiong *Dejing Dou* *Antoni B. Chan*

arXiv 2022

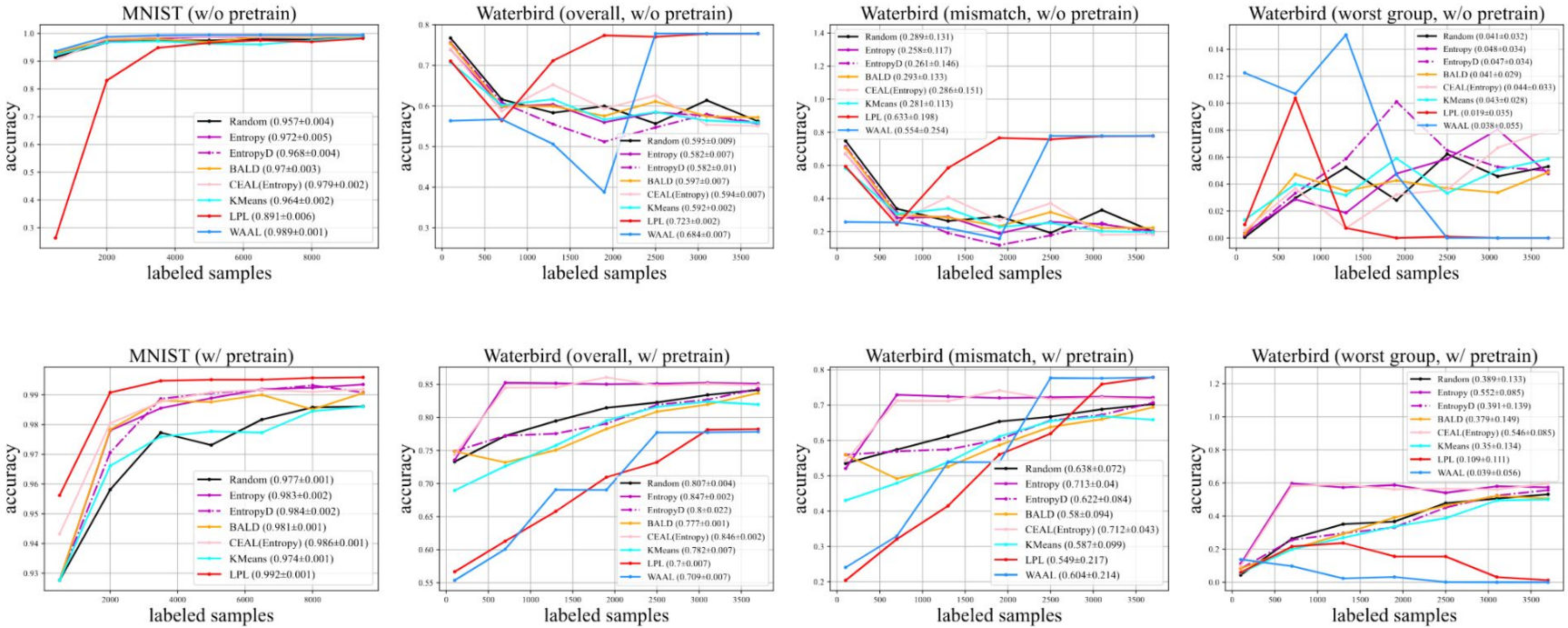


Figure 4: Overall (mismatch, worst group) accuracy vs. budget curves on *MNIST* and *Waterbird* datasets.