



Deep learning based Feature Selection Study

Sanghyuck Lee

2022. 01. 18.

To DO List

➤ Last Week

- Related Works Study and Implementation
- 갑상선 연구 보고 : 지속적 성능 향상 추진 (특징 하위 집합 탐색("슬라이스 단위"))
- 산학협력프로젝트 : 진행 보고

➤ This Week

- Related Works Study and Implementation
- 갑상선 연구 보고 : 지속적 성능 향상 추진 (특징 하위 집합 탐색("세그먼트 단위"))
- 산학협력프로젝트 : 진행 보고

Papers

- Chen, Jianbo, et al. "Learning to explain: An information-theoretic perspective on model interpretation." *International Conference on Machine Learning*. PMLR, 2018.
- Yoon, Jinsung, James Jordon, and Mihaela van der Schaar. "INVASE: Instance-wise variable selection using neural networks." *International Conference on Learning Representations*. 2018.
- Schwab, Patrick, and Walter Karlen. "Cexplain: Causal explanations for model interpretation under uncertainty." *arXiv preprint arXiv:1910.12336* (2019).
- Jethani, Neil, et al. "Have We Learned to Explain?: How Interpretability Methods Can Learn to Encode Predictions in their Interpretations." *International Conference on Artificial Intelligence and Statistics*. PMLR, 2021.
- Panda, Pranoy, Sai Srinivas Kancheti, and Vineeth N. Balasubramanian. "Instance-wise Causal Feature Selection for Model Interpretation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.

Implementation Progress

- INVASE (Completed):
 - Issues: Search proper “Lamda” for Various Datasets
 - Results:

Iteration	Predictor		Baseline		#Features (Total: 64)
	AUC	ACC	AUC	ACC	
1	0.918	0.738	0.917	0.765	64
2	0.919	0.752	0.907	0.734	64
3	0.652	0.192	0.915	0.758	21
4	0.92	0.745	0.914	0.724	64
5	0.922	0.725	0.912	0.77	64

CIFAR-10
(Lamda: 0.05)

Iteration	Predictor		Baseline		#Features (Total: 196)
	AUC	ACC	AUC	ACC	
1	0.943	0.866	0.986	0.899	196
2	0.916	0.814	0.978	0.916	160
3	0.932	0.827	0.969	0.899	167
4	0.929	0.842	0.911	0.815	196
5	0.874	0.599	0.978	0.91	196

Dogs vs Cats
(Lamda: 0.01)

Implementation Progress

- SCM (Completed):
 - Issues: Trade off between Performance and Interpretability
 - Results:

Iteration	Predictor		Baseline		#Features (Total: 64)
	AUC	ACC	AUC	ACC	
1	0.816	0.403	0.95	0.706	30
2	0.749	0.303	0.954	0.72	30
3	0.749	0.3	0.95	0.715	30
4	0.782	0.341	0.951	0.712	30
5	0.8	0.386	0.942	0.695	30

CIFAR-10
(#Features: 30 / 64)

Implementation Progress

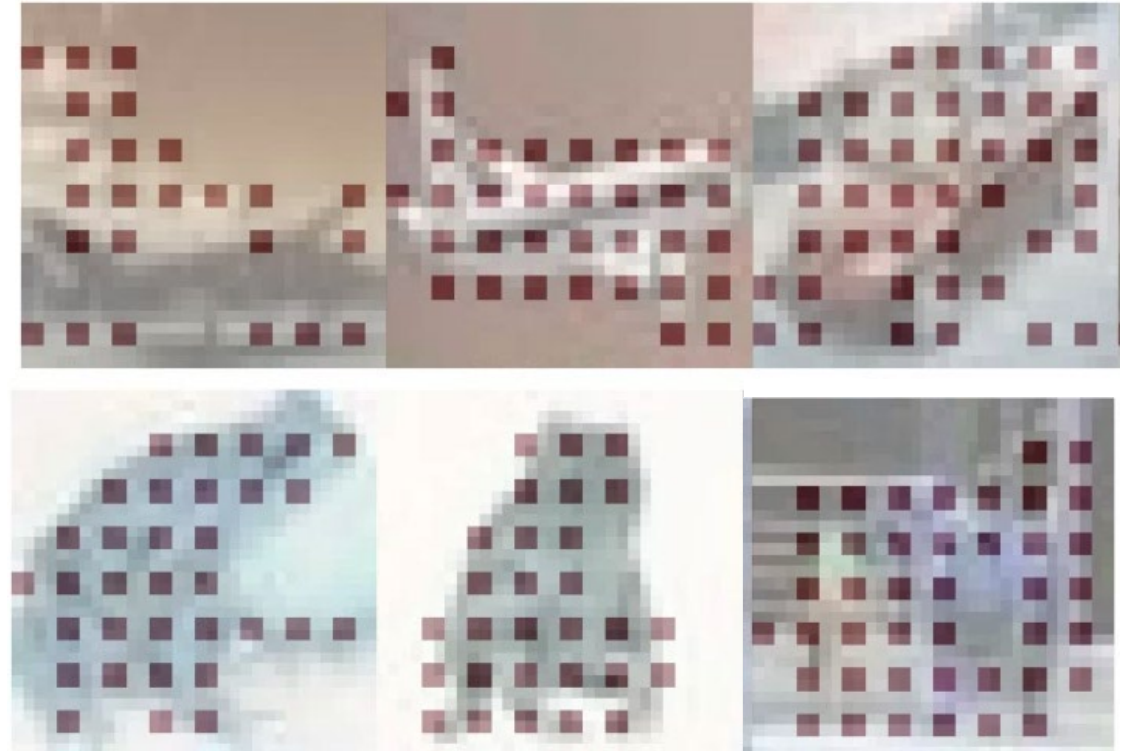
- Real-X (Not Completed):

- Issues:

- Results:

Iteration	Predictor		Baseline		#Features (Total: 256)
	AUC	ACC	AUC	ACC	
1	0.5	0.098	0.914	0.57	66.02
2	0.5	0.101	0.929	0.633	27.34
3	0.5	0.097	0.919	0.615	10.13
4	0.508	0.112	0.929	0.641	0.01
5	0.5	0.102	0.928	0.633	0.24

CIFAR-10
(Lamda: 0.5)



Further Research

- L2X, CXPlain Implements
- Summarizes Advantages and Limitations (모델 압축 관점에서)

방법론	장점	한계점
INVASE	- Prediction 모델을 함께 학습함으로써, 상황이 주어지면 Prediction 성능을 끌어올릴 수도 있음 (이론적으로는 Prediction 성능 유지가 쉬울 수 있음).	- 데이터셋별로 Lamda 값 서치가 필요함에 따라, 효과적인 서치 방식을 찾지 않으면, 성과를 얻어내는데 시간이 오래 걸릴 수 있음. - Tabular 데이터에 대한 방법론만 참고 가능하므로, 진행에 어려움이 있음.
SCM	- 선택 특징 수를 사전에 정의할 수 있음에 따라, 최고 실험 성능을 서치하기가 쉬움(직관적임). - 이미지 데이터셋이 중점인 연구이므로, 추가 연구가 쉬움.	- Prediction 모델과 함께 학습하지 않으므로, Selector-Predictor 성능 유지가 어려울 수 있음. - 해당 성능의 데이터셋 의존성이 클 수 있음.
Real-X	- 시각화 결과를 보아 좋은 모델이 될 잠재력이 있음.	- Lamda 값 서치 필요함. - 성능 아직 잘 안나옴.

갑상선 연구 실험 : 1) AX6 추가 실험(10회 반복)

- 지난주 : AX2, AX3, AX4, LT, RT가 가장 좋은 성능, 아직 AX6 추가시를 고려하지 않음.
- 이번 주 1번 실험 : AX6 추가, AX6 제외 버전을 DenseNet121, 169, 264 버전을 나누어 실험 진행
- 실험 결과 : AUROC 내림차순, AX6를 추가했을 때의 DenseNet121이 가장 뛰어난 성능 출력

Table . Metrics with 95% confidence interval (▼/△ indicates that the corresponding method is **significantly worse** than the best model written in bold based on paired t-test at 5% significance level).

Model	AUROC			AUPRC			Accuracy			F1 Score			Sensitivity TP/(TP+FN)			Specificity TN/(TN+FP)			Precision TP/(TP+FP)		
	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value	Average	95% CI Lower Upper	p-value
AX6 추가 (DenseNet121)	0.754	0.716 0.792		0.459	0.392 0.526		0.709	0.548 0.871	0.598	0.387	0.344 0.430	0.922	0.622	0.471 0.773		0.715	0.508 0.922	0.408	0.335	0.232 0.437	
AX6 제외 (DenseNet169)	0.747	0.695 0.799	0.759	0.422	0.305 0.540	0.552	0.752	0.664 0.840	0.963	0.343	0.237 0.449	0.193	0.523	0.338 0.708	0.078	0.781	0.662 0.899		0.288	0.167 0.409	0.549
AX6 제외 (DenseNet264)	0.745	0.707 0.784	0.639	0.419	0.347 0.491	0.195	0.683	0.537 0.829	0.135	0.341	0.264 0.418	0.059	0.607	0.427 0.787	0.862	0.688	0.497 0.879	0.171	0.314	0.191 0.437	0.791
AX6 추가 (DenseNet264)	0.744	0.686 0.802	0.766	0.41	0.336 0.484	0.378	0.679	0.496 0.862	0.477	0.362	0.316 0.408	0.531	0.614	0.464 0.765	0.908	0.683	0.449 0.918	0.415	0.29	0.222 0.358	0.258
AX6 추가 (DenseNet169)	0.736	0.688 0.783	0.453	0.415	0.315 0.516	0.294	0.745	0.656 0.834	0.804	0.34	0.255 0.426	0.219	0.538	0.368 0.708	0.436	0.775	0.652 0.897	0.94	0.288	0.221 0.354	0.49
AX6 제외 (DenseNet121)	0.724	0.664 0.784	0.218	0.421	0.328 0.513	0.411	0.755	0.633 0.876		0.391	0.324 0.458		0.559	0.437 0.680	0.422	0.777	0.619 0.934	0.958	0.324	0.239 0.408	0.866

CI, confidence interval; *: p-value < 0.05; **: p-value < 0.01

갑상선 연구 실험 : 2) SFS 알고리즘(1회 반복)

- AX2, AX3, AX4, AX6, LT, RT(6개 슬라이스)의 총 58개 특징을 사용한 SFS 알고리즘 구현
 - 1) 각 58개 특징을 DenseNet121에 입력하여 특징 Evaluation을 진행하고, 랭킹을 매김.
 - 2) 랭킹을 기반으로 1위부터 58위까지 점차적으로 입력 특징 하위 집합에 포함하여 57개의 Chromosome에 대한 성능 평가를 진행함.
 - 3) 따라서, 1개 특징을 사용한 58번의 성능 평가와 2개 이상의 특징을 사용한 57번의 성능 평가가 진행됨. ($58+57=106$ FFCs)
- * 이러한 Chromosome을 위해, 매 Iteration마다, 각 후보 특징 집합의 Evaluation을 진행할 경우, 1회 실험당 약 12일 정도의 시간이 소요되어, 위와 같이, 첫번째 단계에서만 특징 Evaluation을 진행하여 SFS를 수행함(약 15 시간).

갑상선 연구 실험 : 2) SFS 알고리즘(1회 반복)

• 실험 결과 : AUROC 내림차순

Rank	Segment	Score
1	AX2 외직근 left	0.835
2	AX2 안와 right	0.826
3	CO2 하직근 left	0.813
4	AX2 내직근 left	0.810
5	RT 상직근	0.807
6	AX3 내직근 right	0.804
7	AX3 안와 right	0.798
8	RT 하직근	0.797
9	CO2 외직근 left	0.796
10	AX2 내직근 right	0.794
⋮	⋮	⋮
58	AX3 눈 right	0.449

Rank	#Features	AUROC	AUPRC	Accuracy	Macro F1	Sensitivity	Specificity	Precision	Time
1	2	0.917	0.517	0.615	0.311	1.000	0.578	0.184	0:03:38
2	37	0.884	0.549	0.907	0.483	0.500	0.946	0.467	0:13:46
3	24	0.867	0.497	0.888	0.550	0.786	0.898	0.423	0:09:01
4	57	0.866	0.551	0.801	0.385	0.714	0.810	0.263	0:22:38
5	3	0.866	0.388	0.814	0.375	0.643	0.830	0.265	0:03:48
6	21	0.864	0.665	0.888	0.550	0.786	0.898	0.423	0:07:57
7	50	0.860	0.323	0.863	0.389	0.500	0.898	0.318	0:19:20
8	23	0.856	0.568	0.907	0.516	0.571	0.939	0.471	0:08:37
9	32	0.855	0.577	0.888	0.500	0.643	0.912	0.409	0:11:52
10	51	0.855	0.525	0.863	0.421	0.571	0.891	0.333	0:19:41
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
15	58 (모두 사용)	0.836	0.544	0.926	0.400	0.286	0.986	0.667	0:23:12
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
117	1	0.449	0.076	0.714	0.115	0.214	0.762	0.079	0:03:20

추후 진행 방향

- SFS 알고리즘 반복 실험
- 이정규 교수님 미팅 : 내일 모레 (01.20 목)
 - 현재 실험 보고
 - 모델 하나를 골라 Class Activation Map 보고

산학프로젝트

- **SW 등록 3건(1~2월)**

- 기계 데이터를 위한 멀티모달 학습 연구 : 아성이 SW 제작, 태원이 등록 실무
- 보안 텍스트 데이터 분석을 위한 딥러닝 기법 연구 : 건일이 SW 제작, 수정이 등록 실무
- 임베디드 AI 기반 Scene Text Recognition 연구 : 건일이 SW 제작, 수정이 등록 실무

- **국제 학술대회 논문 2편 + 국내 특허출원 2건(6~8월)**

- 취향에 맞는 음악 추천을 위한 음악 장르 분류 프로젝트 : 아성+성현 KCI 논문 후속작(논문 1편 + 특허 1건)
- 의료 영상 분석을 위한 XAI 연구 : 상혁 졸업 논문 후속작(논문 1편 + 특허 1건)