



Deep learning based Feature Selection Study

Sanghyuck Lee

2022. 02. 22.

To DO List

➤ Last Week

- Related Works Study and Implementation
- 갑상선 연구 보고 : W/O meta data, With meta data 실험 결과 보고
- 산학협력프로젝트 : 진행 보고

➤ This Week

- Related Works Study and Implementation
- 갑상선 연구 보고 : Class Activation Map 보고
- 졸업 논문 작업(신규)
- 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 (Completed):
 - Issues: Validation with SOTA models
 - Results:

Iteration	Predictor		Baseline		#Features (Total: 256)
	AUC	ACC	AUC	ACC	
1	0.940	0.646	0.926	0.619	129.14
2	0.930	0.641	0.936	0.654	137.30
3	0.941	0.645	0.933	0.635	146.32

**CIFAR-10
(Lamda: 0.5)**



Implementation Progress

- CXPlain (Not Completed):
 - Issues: Many Hyper Parameters,
 - Results:

Iteration	Predictor		Baseline		#Features (Total: 256)
	AUC	ACC	AUC	ACC	
1	0.546	0.154	0.615	0.293	128

CIFAR-10
(#Features: 128 / 256)

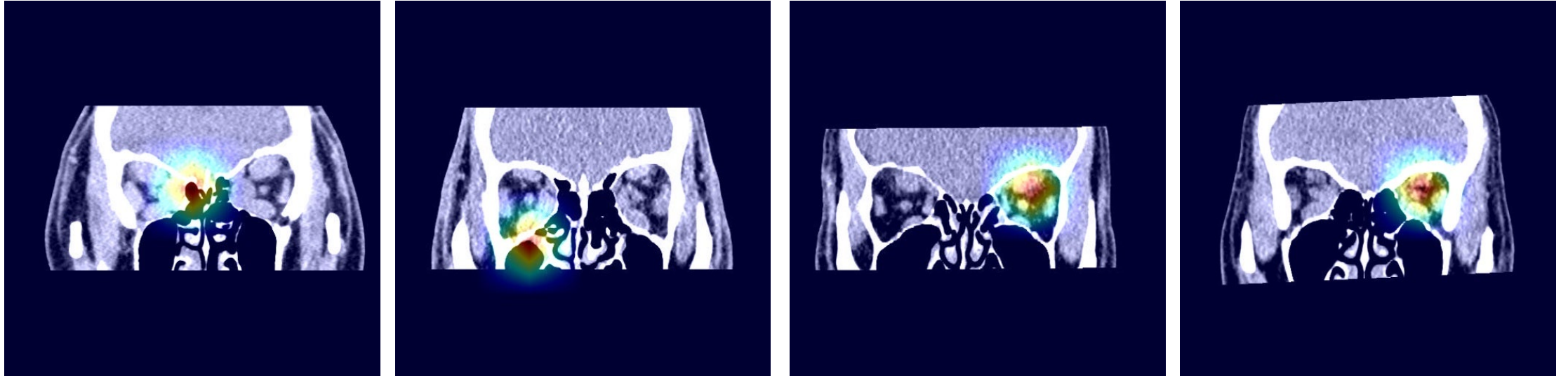
Further Research

- L2X 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	- 이미지 데이터셋이 중점인 연구이므로, 추가 연구가 쉬움. - INVASE 보다 Lamda 값 서치가 쉬울 것으로 생각됨.	- Lamda 값 서치 필요함.
CXPlain		- Hyper Parameter가 비교적 많은 것으로 보임.

갑상선 연구 : Class Activation Map (CO2)

- 기본적인 CAM 기법으로는 현재 모델에 적용이 어려운 것 같아, CO2만을 입력한 모델에 대해서 CAM 이미지를 출력함.



졸업 논문 작업

- 주제 :

의료 영상 분석을 위한 특징 선택 기반 설명 가능한 인공지능 프레임워크

Explainable Artificial Intelligence based on Feature Selection for Medical Image Analysis

- 추가할 실험 내용 :

현재까지 진행한 SFS, UFS, Total Feature 실험에, No 세그멘테이션 실험을 추가함.

=> 세그멘테이션과 Feature Selection을 결합하여, 진단 성능과 설명력 모두를 확보함.

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)