



Deep learning based Feature Selection Study

Sanghyuck Lee

2022. 02. 08.

To DO List

➤ Last Week

- Related Works Study and Implementation
- 갑상선 연구 보고 : SFS를 통한 Best 슬라이스 Subset 탐색
- 산학협력프로젝트 : 진행 보고

➤ This Week

- Related Works Study and Implementation
- 갑상선 연구 보고 : 1회 실험이 아닌, 10회 실험 평균을 통해 Ranking을 매기는 접근 방식을 실험
- 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:

- Results:

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	- 이미지 데이터셋이 중점인 연구이므로, 추가 연구가 쉬움. - INVASE 보다 Lamda 값 서치가 쉬울 것으로 생각됨.	- Lamda 값 서치 필요함.

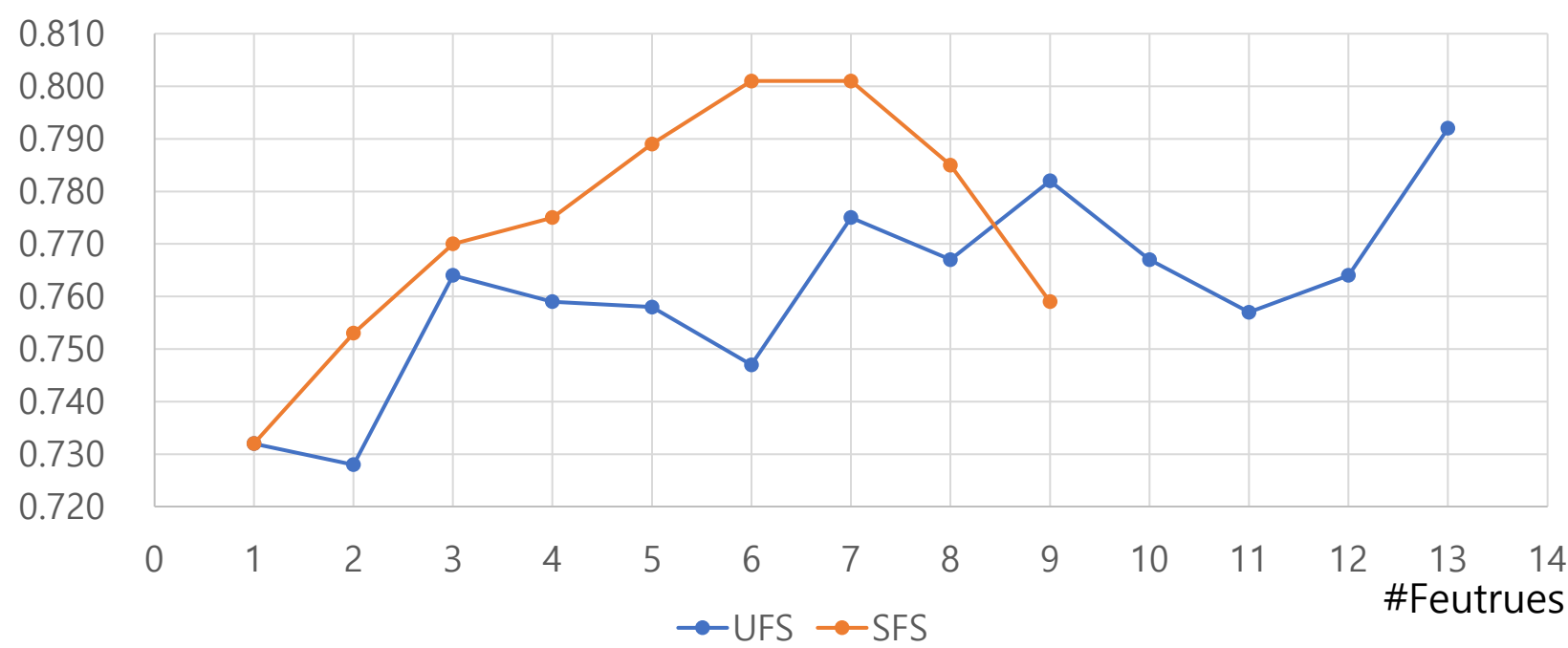
갑상선 연구

Table 3. 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).

Subset	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
SFS	0.801	0.768 0.834		0.461	0.395 0.528		0.741	0.629 0.853	0.501	0.434	0.379 0.489		0.704	0.592 0.816		0.742	0.595 0.888	0.271	0.337	0.251 0.424	0.442
UFS (Human)	0.792	0.756 0.829	0.61	0.453	0.392 0.515	0.857	0.768	0.655 0.881		0.347	0.300 0.395	0.018	0.48	0.302 0.659	0.011	0.803	0.649 0.957		0.403	0.235 0.570	

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

AUROC



Order	UFS	SFS
1	CO2	CO2
2	Age	AX5
3	SA (RT)	Age
4	AX3	AX3
5	AX5	AX6
6	AX2	Sex
7	CO3	AX2
8	SA (LT)	SA (RT)
9	AX4	AX1
10	AX1	
11	CO1	
12	Sex	
13	AX6	

MAX
AUROC

MAX
AUROC

갑상선 연구

- 사진 데이터 수신



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)