



Deep learning based Feature Selection Study

Sanghyuck Lee

2022. 01. 11.

To DO List

➤ Last Week

- Related Works Study and Implementation
- 갑상선 연구 보고 : Coronal -> Axial 등의 순서로 실험
- 산학협력프로젝트 : 계획 보고

➤ This Week

- Related Works Study and Implementation
- 갑상선 연구 보고 : 지속적 성능 향상 추진 (입력 데이터 조합 탐색)
- 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: Search proper “Lamda” for Various Datasets
 - 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)

Further Research

- Real-X, 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.

갑상선 연구 : 10회 반복 (AUPRC 기준 내림차순)

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
co2	0.706	0.653 0.760	0.106	0.381	0.316 0.445		0.78	0.696 0.863		0.362	0.291 0.433		0.476	0.347 0.605	0.003 (**▼)	0.825	0.715 0.934		0.351	0.240 0.461	
ax2, ax3, ax4	0.734	0.684 0.784		0.375	0.272 0.477	0.867	0.718	0.615 0.821	0.293	0.342	0.272 0.412	0.515	0.55	0.405 0.695	0.027	0.737	0.604 0.871	0.246	0.269	0.186 0.352	0.035 (*▼)
ax1, ax2, ax3, ax4, ax5	0.715	0.657 0.773	0.42	0.361	0.257 0.465	0.676	0.732	0.572 0.891	0.505	0.331	0.261 0.400	0.285	0.496	0.310 0.683	0.016 (*▼)	0.764	0.562 0.966	0.494	0.308	0.242 0.375	0.386
lt, rt	0.686	0.589 0.783	0.218	0.361	0.243 0.480	0.651	0.569	0.439 0.699	0.011	0.314	0.230 0.398	0.117	0.691	0.589 0.793	0.241	0.553	0.399 0.708	0.008 (**▼)	0.22	0.137 0.304	0.009 (**▼)
rt	0.723	0.667 0.778	0.743	0.342	0.260 0.424	0.371	0.601	0.460 0.742	0.009 (**▼)	0.327	0.281 0.373	0.304	0.708	0.540 0.876	0.286	0.579	0.393 0.765	0.008 (**▼)	0.244	0.173 0.315	0.078
co2, co3	0.718	0.673 0.763	0.454	0.33	0.286 0.373	0.126	0.701	0.599 0.802	0.203	0.335	0.297 0.373	0.318	0.586	0.426 0.746	0.066	0.718	0.581 0.856	0.187	0.272	0.202 0.342	0.092
co1, co2, co3	0.689	0.640 0.739	0.087	0.323	0.262 0.384	0.184	0.747	0.653 0.840	0.562	0.332	0.279 0.385	0.441	0.489	0.350 0.628	0.004 (**▼)	0.783	0.656 0.910	0.559	0.303	0.211 0.396	0.416
co1, co2	0.663	0.604 0.723	0.039	0.321	0.246 0.397	0.262	0.533	0.319 0.747	0.04 (*▼)	0.288	0.234 0.342	0.1	0.627	0.438 0.817	0.081	0.511	0.238 0.785	0.042 (*▼)	0.234	0.139 0.329	0.153
ax3	0.689	0.632 0.746	0.117	0.309	0.211 0.408	0.089	0.648	0.463 0.832	0.141	0.305	0.206 0.404	0.253	0.586	0.363 0.810	0.128	0.649	0.405 0.894	0.137	0.247	0.124 0.371	0.059
lt	0.666	0.581 0.751	0.204	0.305	0.219 0.391	0.125	0.506	0.323 0.689	0.007 (**▼)	0.292	0.244 0.339	0.099	0.728	0.554 0.902	0.502	0.469	0.237 0.702	0.007 (**▼)	0.218	0.135 0.300	0.042 (*▼)
ax5	0.676	0.617 0.734	0.104	0.285	0.219 0.351	0.031 (*▼)	0.706	0.588 0.823	0.22	0.334	0.292 0.375	0.432	0.545	0.398 0.691	0.013 (*▼)	0.723	0.569 0.876	0.18	0.284	0.195 0.373	0.268
co1, co3	0.648	0.603 0.692	0.017	0.271	0.214 0.327	0.008 (**▼)	0.668	0.478 0.858	0.255	0.209	0.117 0.301	0.021	0.429	0.154 0.703	0.026 (*▼)	0.695	0.435 0.954	0.31	0.196	0.121 0.271	0.063
ax4	0.656	0.576 0.737	0.113	0.263	0.178 0.347	0.034 (*▼)	0.698	0.539 0.856	0.33	0.299	0.235 0.362	0.139	0.479	0.279 0.679	0.012 (*▼)	0.723	0.519 0.927	0.342	0.271	0.188 0.354	0.2
ax1	0.591	0.527 0.656	0.001 (**▼)	0.226	0.166 0.286	0.001 (**▼)	0.588	0.375 0.800	0.082	0.231	0.170 0.292	0.004 (**▼)	0.505	0.238 0.773	0.051	0.59	0.306 0.874	0.1	0.234	0.101 0.367	0.15
ax2	0.615	0.531 0.699	0.05	0.223	0.158 0.288	0.008 (**▼)	0.591	0.412 0.770	0.076	0.284	0.241 0.328	0.082	0.587	0.385 0.789	0.048	0.584	0.355 0.812	0.07	0.231	0.148 0.314	0.144
co3	0.577	0.521 0.633	0.003 (**▼)	0.219	0.157 0.280	0.004 (**▼)	0.479	0.265 0.692	0.013 (*▼)	0.215	0.155 0.275	0.008 (**▼)	0.62	0.348 0.891	0.19	0.45	0.164 0.737	0.018 (*▼)	0.149	0.090 0.208	0.011 (*▼)
co1	0.562	0.527 0.596	0 (**▼)	0.152	0.123 0.181	0 (**▼)	0.355	0.201 0.508	0.001 (**▼)	0.245	0.205 0.286	0.005 (**▼)	0.791	0.671 0.911		0.288	0.097 0.480	0.001 (**▼)	0.149	0.118 0.180	0.001 (**▼)

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

갑상선 연구 : 10회 반복 (AUPRC 기준 내림차순)

- 금주 실험 1 : (co2)를 고정한 채로, (ax3), (ax2~ax4), (**ax1~ax5**), (lt), (rt), (lt, rt)에 대한 실험

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
co2, ax1, ax2, ax3, ax4, ax5	0.725	0.644 0.807	0.835	0.391	0.304 0.478		0.714	0.549 0.879	0.148	0.352	0.273 0.431	0.702	0.529	0.350 0.708	0.709	0.738	0.522 0.953	0.139	0.354	0.223 0.486	0.625
co2 (prev)	0.706	0.653 0.760	0.106	0.381	0.316 0.445	0.806	0.78	0.696 0.863	0.667	0.362	0.291 0.433		0.476	0.347 0.605	0.342	0.825	0.715 0.934	0.66	0.351	0.240 0.461	0.555
co2, lt, rt	0.705	0.628 0.782	0.434	0.377	0.286 0.468	0.838	0.727	0.569 0.885	0.335	0.331	0.265 0.396	0.525	0.511	0.330 0.692	0.634	0.753	0.551 0.954	0.34	0.324	0.226 0.423	0.308
ax2, ax3, ax4 (prev)	0.734	0.684 0.784		0.375	0.272 0.477	0.746	0.718	0.615 0.821	0.119	0.342	0.272 0.412	0.515	0.55	0.405 0.695		0.737	0.604 0.871	0.102	0.269	0.186 0.352	0.069
co2, lt	0.691	0.627 0.755	0.28	0.373	0.290 0.455	0.761	0.734	0.596 0.872	0.016	0.345	0.267 0.424	0.703	0.517	0.344 0.691	0.721	0.761	0.578 0.943	0.015	0.295	0.201 0.390	0.133
co2, ax3	0.709	0.673 0.745	0.224	0.367	0.298 0.437	0.456	0.806	0.697 0.915		0.358	0.300 0.417	0.919	0.403	0.258 0.548	0.044	0.859	0.713 1.004		0.395	0.297 0.493	
ax1, ax2, ax3, ax4, ax5 (prev)	0.715	0.657 0.773	0.42	0.361	0.257 0.465	0.443	0.732	0.572 0.891	0.386	0.331	0.261 0.400	0.285	0.496	0.310 0.683	0.544	0.764	0.562 0.966	0.4	0.308	0.242 0.375	0.14
co2, rt	0.686	0.611 0.761	0.289	0.358	0.275 0.440	0.451	0.737	0.623 0.851	0.128	0.339	0.257 0.421	0.637	0.506	0.369 0.643	0.465	0.766	0.626 0.906	0.099	0.37	0.189 0.551	0.79
co2, ax2, ax3, ax4	0.718	0.654 0.781	0.418	0.345	0.274 0.416	0.135	0.739	0.663 0.815	0.086	0.317	0.241 0.393	0.139	0.52	0.330 0.710	0.689	0.766	0.661 0.870	0.087	0.26	0.217 0.303	0.017

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

갑상선 연구 : 10회 반복 (AUPRC 기준 내림차순)

- 금주 실험 2 : Best 모델에 대해 (lt, rt)를 추가하여 실험 진행

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
co2, ax1, ax2, ax3, ax4, ax5 (prev)	0.725	0.644 0.807		0.391	0.304 0.478		0.714	0.549 0.879	0.503	0.352	0.273 0.431		0.529	0.350 0.708		0.738	0.522 0.953	0.535	0.354	0.223 0.486	
co2, ax1, ax2, ax3, ax4, ax5, lt, rt	0.723	0.669 0.778	0.914	0.386	0.306 0.465	0.836	0.743	0.639 0.848		0.349	0.294 0.403	0.908	0.526	0.367 0.685	0.962	0.775	0.638 0.911		0.33	0.225 0.434	0.7

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

갑상선 연구 : 10회 반복 (AUPRC 기준 내림차순)

- 금주 실험 3 : lt, rt 정보 활용을 위해, ax 정보를 줄임. -> (ax1~ax5) -> **(ax2~ax4)** -> (ax3)

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
co2, ax2, ax3, ax4, lt, rt	0.745	0.707 0.784		0.419	0.347 0.491		0.683	0.537 0.829	0.117	0.341	0.264 0.418	0.377	0.607	0.427 0.787		0.688	0.497 0.879	0.105	0.314	0.191 0.437	0.254
co2, ax3, lt, rt	0.736	0.695 0.778	0.706	0.416	0.344 0.487	0.931	0.806	0.747 0.864		0.384	0.309 0.460		0.49	0.333 0.647	0.225	0.853	0.772 0.934		0.383	0.287 0.479	
co2, ax1, ax2, ax3, ax4, ax5 (prev)	0.725	0.644 0.807	0.562	0.391	0.304 0.478	0.542	0.714	0.549 0.879	0.201	0.352	0.273 0.431	0.544	0.529	0.350 0.708	0.383	0.738	0.522 0.953	0.197	0.354	0.223 0.486	0.691
co2, ax1, ax2, ax3, ax4, ax5, lt, rt	0.723	0.669 0.778	0.409	0.386	0.306 0.465	0.39	0.743	0.639 0.848	0.077	0.349	0.294 0.403	0.454	0.526	0.367 0.685	0.484	0.775	0.638 0.911	0.061	0.33	0.225 0.434	0.304

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

갑상선 연구 : 10회 반복 (AUPRC 기준 내림차순)

- 금주 실험 4 : Best에서, Parameter Size 변경 실험

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
DenseNet169	0.747	0.695 0.799		0.422	0.305 0.540		0.752	0.664 0.840	0.963	0.343	0.237 0.449	0.193	0.523	0.338 0.708	0.242	0.781	0.662 0.899		0.288	0.167 0.409	0.278
DenseNet121	0.724	0.664 0.784	0.24	0.421	0.328 0.513	0.978	0.755	0.633 0.876		0.391	0.324 0.458		0.559	0.437 0.680	0.398	0.777	0.619 0.934	0.958	0.324	0.239 0.408	
DenseNet264	0.745	0.707 0.784	0.908	0.419	0.347 0.491	0.944	0.683	0.537 0.829	0.135	0.341	0.264 0.418	0.059	0.607	0.427 0.787		0.688	0.497 0.879	0.171	0.314	0.191 0.437	0.859

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

추후 진행 방향

- 모델 구조 개선

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)