Deep learning based Feature Selection Study

Sanghyuck Lee

2021. 09 14.

Outline

- INVASE: INSTANCE-WISE VARIABLE SELECTION USING NEURAL NETWORKS
- 중견 연구 과제
- 갑상선 연구 과제
- 산학 협력 과제
- 11월 특강 자료 작성

Papers for Instance-Wise Feature Selection

Yoon, Jinsung, James Jordon, and Mihaela van der Schaar. "INVASE: Instance-wise variable selection using neural networks." *International Conference on Learning Representations*. 2018.

=> ICLR

Tonekaboni, Sana, et al. "What went wrong and when? Instance-wise feature importance for time-series black-box models." *Advances in Neural Information Processing Systems* 33 (2020).

=> NeurIPS

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.

=> CVPR

Liyanage, Yasitha Warahena, and Daphney–Stavroula Zois. "Optimum Feature Ordering for Dynamic Instance–Wise Joint Feature Selection and Classification." *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021.

=> ICASSP

Paper Review

Paper Title: INVASE: Instance-wise variable selection using neural networks

Authors: Jinsung Yoon, James Jordon, Mihaela van der Schaar

Conference: ICLR 2018

Abstract of the Paper

The advent of big data brings with it data with more and more dimensions and thus a growing need to be able to efficiently select which features to use for a variety of problems. While global feature selection has been a well-studied problem for quite some time, only recently has the paradigm of instance-wise feature selection been developed. In this paper, we propose a new **instance-wise feature selection method**, which we term INVASE. INVASE consists of **3 neural networks, a selector network, a predictor network and a baseline network** which are used to **train the selector** network using the actor-critic methodology. Using this methodology, INVASE is capable of **flexibly discovering feature subsets of a different size for each instance**, which is a **key limitation** of existing state-of-the-art methods. We demonstrate through a mixture of synthetic and real data experiments that INVASE significantly outperforms state-of-the-art benchmarks.

Introduction of the Paper

High-dimensional data is becoming more readily available, and it brings with it a growing **need to be able to efficiently select which features** to use for a variety of problems. When doing predictions, it is well known that using **too many variables with too few samples can lead to overfitting**, which can significantly hinder the performance of predictive models. In the realm of **interpretability**, the **large dimensionality** of the data is often too much information to present to a human who may be using the machine learning model as a support system. **Understanding** which **features** are most **relevant** to an outcome or to a model **output** is an important first step in improving **predictions** and **interpretability** and many works exist that tackle feature selection on a global level. However, in the **heterogeneous data** we typically encounter, the prediction made by a model (and indeed the true label) may rely on a different subset of the features for different subgroups within the data [14]. In this paper we propose a novel **instance-wise feature selection** method, INVASE (INstance-wise VAriable SElection), which attempts to learn which subset of the features is relevant for each sample, allowing us to display the minimal information required to explain each prediction and also to **reduce overfitting** of predictive models.

Introduction of the Paper

Discovering a global subset of relevant features for a particular task is a well-studied problem and there are several existing methods for solving it such as Sequential Correlation Feature Selection [11], Mutual Information Feature Selection [21], Knockoff models [3], and more [10; 16]. However, **global feature selection** suffers from a key limitation - the features discovered by global feature selection are the same for all samples. In many cases, in particular when populations are highly heterogeneous, the relevant features may differ across samples [33; 32]. For instance, different patient subgroups have different relevant features for predicting heart failure [14]. Instance-wise feature selection methods such as [4; 27] instead try to discover the features that are relevant for each sample. When the goal is to **provide an interpretable explanation** of the predictions made, a key challenge is in **ensuring that we do not over-explain by providing too much information** (i.e. **choosing too many features**). Naturally, by performing feature selection on an individualized level we are able to select features that are more relevant to each sample, rather than having to choose the top k features globally, which may not explain the predictions for some samples very well, but simply perform well on average across all samples.

Introduction of the Paper

In this paper, we propose a novel instance-wise feature selection method which we term INVASE. We draw influence from actor-critic models [22] to solve the problem of backpropagating through subset sampling. Our model consists of 3 neural networks: a selector network, a predictor network and a baseline network. During training, each of these are trained iteratively, with the selector network being trained to minimize a Kullback-Leibler (KL) divergence between the full conditional distribution and the selected-features-only conditional distribution of the outcome. Our model is **capable of discovering a different number of relevant variables for each sample** which is a **key limitation** in **existing instance-wise** approaches (such as [4]). We show significant improvements over the state-of-the-art in both **synthetic data** and **real-word data** in terms of **true positive rates**, **false discovery rates**, and show better **predictive performance** with respect to several prediction metrics. Our model can also be easily extended to handle both continuous and discrete outputs and time-series inputs (see the Appendix for details).

Experiments of the Paper (Settings)

In this section, we quantitatively evaluate INVASE against various state-of-the-art benchmarks on both synthetic and real-world datasets. We evaluate our performance both at identifying ground truth relevance and at enhancing predictions. We compare our model with **4 global variable selection models**: Knockoffs [3], Tree Ensembles (Tree) [7], Sequential Correlation Feature Selection (SCFS) [11], and LASSO regularized linear model; and **3 instance-wise feature selection methods**: L2X [4], LIME [24], and Shapley [18]. The details of benchmark implementation can be found in the appendix. Implementation of INVASE can be found at https://github.com/jsyoon0823/ INVASE.

Experiments of the Paper (Synthetic)

4.1.2 DISCOVERY

Note that in Syn4 and Syn5, the number of relevant features is different for different samples.

Dataset	Syn1		Syn2		Syn3		Syn4		Syn5		Syn6	
Metrics (%)	TPR	FDR	TPR	FDR	TPR	FDR	TPR	FDR	TPR	FDR	TPR	FDR
INVASE	**100.0**	**0.0**	**100.0**	**0.0**	**92.0**	**0.0**	**99.8**	**10.3**	**84.8**	**1.1**	**90.1**	**7.4**
L2X	100.0	0.0	100.0	0.0	69.4	30.6	79.5	21.8	74.8	26.3	83.3	16.7
LIME	13.8	86.2	100.0	0.0	98.1	1.9	40.7	49.4	41.1	50.6	50.5	49.5
Shapley	60.4	39.6	93.3	6.7	90.9	9.1	65.2	31.9	62.9	33.7	71.2	28.8
Knockoff	10.0	70.0	8.7	36.2	81.2	17.5	38.8	35.1	41.0	51.1	56.6	42.1
Tree	100.0	0.0	100.0	0.0	100.0	0.0	54.7	39.0	56.8	37.5	60.0	40.0
SCFS	23.5	76.5	39.5	60.5	78.3	22.0	48.9	52.4	42.4	51.2	56.1	43.9
LASSO	19.0	81.0	39.8	60.2	78.3	21.7	49.9	50.9	45.5	48.2	56.4	43.6

Experiments of the Paper (Synthetic)

4.1.3 PREDICTION

In this experiment we analyze the effect of using feature selection as **a pre-processing step for prediction**. We first **perform feature selection** (either instance-wise or global) and then **train a 3-layer fully connected network** with Batch Normalization [12] in every layer (to avoid overfitting) to **perform predictions** on top of the (feature-selected) data.

Experiments of the Paper (Synthetic)

4.1.3 PREDICTION

Dataset	AUROC					
	w/o FS	with Global	**with INVASE**	with Tree	with L2X	with LASSO
Syn1	.578±.004	.686±.005	.690±.006	.574±.101	.498±.005	.498±.006
Syn2	.789±.003	.873±.003	.877±.003	.872±.003	.823±.029	.555±.061
Syn3	.854±.004	.900±.003	.902±.003	.899±.001	.862±.009	.886±.003
Syn4	.558±.021	.774±.006	.787±.004	.684±.017	.678±.024	.514±.031
Syn5	.662±.013	.784±.005	.784±.005	.741±.004	.709±.008	.691±.024
Syn6	.692±.015	.858±.004	.877±.003	.771±.031	.827±.017	.727±.025
Dataset	**AUPRC**					
	w/o FS	with Global	**with INVASE**	with Tree	with L2X	with LASSO
Syn1	.567±.007	.690±.006	.694±.006	.577±.102	.498±.007	.499±.008
Syn2	.799±.005	.878±.005	.886±.004	.878±.004	.817±.031	.591±.037
Syn3	.861±.003	.905±.002	.907±.003	.904±.002	.860±.012	.890±.002
Syn4	.572±.019	.794±.006	.804±.004	.681±.031	.672±.025	.536±.025
Syn5	.665±.019	.796±.005	.797±.006	.765±.003	.719±.011	.680±.040
Syn6	.709±.018	.870±.005	.886±.004	.779±.027	.835±.017	.757±.036

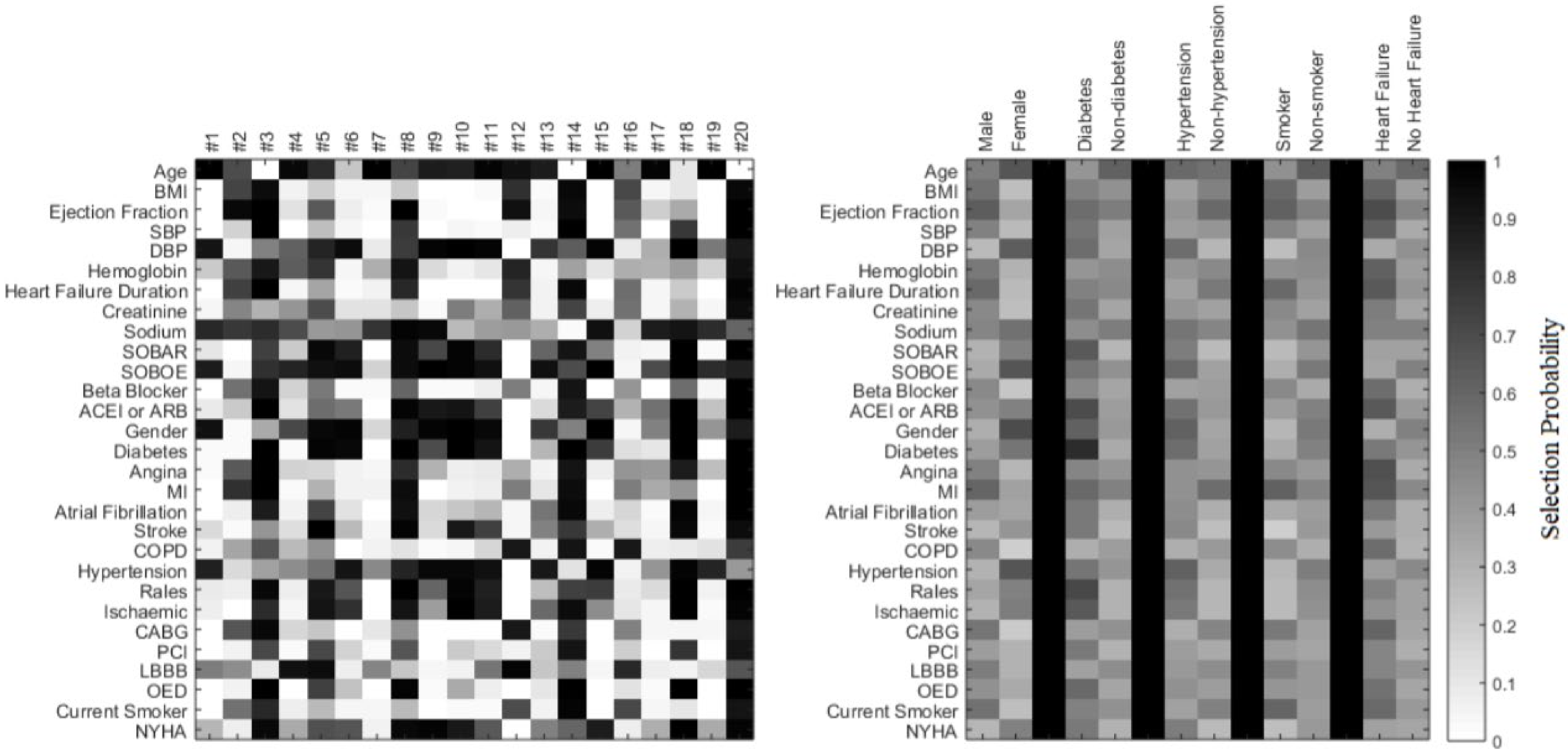
Experiments of the Paper (Real)

4.2.1 DATA DESCRIPTION

The first, the Meta-Analysis Global Group in Chronic Heart Failure (MAGGIC) dataset [23], has 40,409 patients each with **31 measured features**. The label is all-cause **mortality**.

The second, the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial in the US and the European Randomized Study of Screening for Prostate Cancer (ERSPC) dataset [8; 26] contains 38,001 each with **106 measured features**. The label in this dataset is **mortality** due to prostate cancer.

Experiments of the Paper (Real)



Experiments of the Paper (Real)

4.2.3 RESULTS: PREDICTION USING REAL DATA VARIABLES WITH REAL LABEL

Evaluating the performance of feature selection methods on real data is **difficult**, since **ground truth relevance is often not known**. We therefore cannot use TPR and FDR to evaluate the performance on real data. In our final experiment, therefore, we instead **focus on prediction performance** exactly as in 4.1.3 (except now both the features and label come from real data)

Datasets	Metrics	AUROC	AUPRC	AUROC	AUPRC
MAGGIC	Labels	**3 year**		**5 year**	
	INVASE	**.722±.005**	**.655±.010**	**.740±.005**	**.867±.006**
	Without INVASE	.720±.006	.639±.009	.730±.006	.855±.004
PLCO	Labels	**5 year**		**10 year**	
	INVASE	**.637±.007**	**.329±.013**	**.673±.007**	**.506±.006**
	Without INVASE	.629±.008	.324±.011	.657±.006	.485±.008

Summary of the Paper

- Instance-wise Feature Selection
- Unlike the existing Instance-wise FS method, select a different number of features for each instance.
- Experimentally, select different features for each instance in deed, select ground truth features and improve predictive performance.

\+ using 3 networks (selector, predictor, baseline)

In terms of Accuracy and Time

- 정확도 관점:

  정확도가 떨어지면서, 해석 가능성이 높아지는 것은 의미없는 것.

  왜냐하면, 해석이 가능해지지만, 잘못된 해석을 하게 됨.
- 시간 관점:

  해석 가능성: 정의하기 힘든 개념, 특징의 수로 정의

  해석 가능성이 증가하면(특징의 수가 줄면), 시간이 빨라짐.

정리

해석 가능성 -> 정확도가 유지되어야 의미 있음.

해석 가능성 -> 정의하기 힘든 개념, 시간/특징 선택과 여전히 혼동중.

중견 연구 제안서 작성

- 이번 보고 메일과 함께 초안 회신

갑상선 연구

- 중앙대병원측 세그멘테이션 검토 완료 후, Activity 실험 계획

산학 협력 프로젝트

- 특허 출원: 출원 지시 수행 완료

11월 특강 자료 작성

- 주제: ‘사회적 약자와 보편적 복지 향상을 위해 활용될 수 있는 AI 기술 조사’

선별주의는 **자산조사**에 의해 복지기본선 이하의 소외계층, 즉 국민기초생활보장제도상 선정대상을 주된 복지정책의 대상으로 하자는 관점

보편주의는 사회적 시민권을 지반으로 **자산조사에 관계**없이 누구에게도 차별을 제도화하지 않으며 모든 국민을 대상으로 사회복지정책 지원을 하자는 관점 (손병덕, 2020)

1. 사회적 약자를 위해 활용될 수 있는 AI 기술:
    1) AI 로봇, AI 스피커: 독거노인 등 취약 계층을 위한
    2) 교육 AI: 외국인 및 소년 소녀 가장을 위한
    3) 장애인 생활 보조 AI

2. 보편적 복지 향상을 위해 활용될 수 있는 AI 기술:
    1) 아동학대 방지를 위한 AI 시스템
    2) AI 기반 고령 친화 스마트 시티

초안 보고 예정일: ~~9/21 화~~, 9/28 화

References

손병덕, “사회복지정책론”, 학지사, 2020.

Liyanage, Yasitha Warahena, and Daphney–Stavroula Zois. "Optimum Feature Ordering for Dynamic Instance–Wise Joint Feature Selection and Classification." *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 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.

Tonekaboni, Sana, et al. "What went wrong and when? Instance-wise feature importance for time-series black-box models." *Advances in Neural Information Processing Systems* 33 (2020).

Yoon, Jinsung, James Jordon, and Mihaela van der Schaar. "INVASE: Instance-wise variable selection using neural networks." *International Conference on Learning Representations*. 2018.