ARTIFICIAL INTELLIGENCE FOR DETERMINING THE PRESENCE OF HEART DISEASE BY ECG ON THE PTB-XL DATASET

RAZIN VYACHESLAV VASILIEVICH

Young Researcher Laboratory, Institute of Information Technologies, Mathematics and Mechanics, Lobachevsky University, Nizhny Novgorod, Russia
razinvyacheslav@unn.ru

INTRODUCTION

Determining whether a person has heart problems by ECG recording is an extremely relevant and important task in medicine. The paper explores the use of deep learning to solve this problem. PTB-XL is one of the world's largest open heart disease datasets hosted by PhysioNet. The problem of binary classification of the presence of heart diseases or their absence according to the ECG records from this dataset is solved. Approaches for the classification of time series are described in many works using various architectures of neural networks and methods.

MATERIALS & METHODS

The problem of multilabel classification of predicting the main classes of the PTBXL set is solved. The quality metric is ABS (accuracy on a set of classes).The input data for neural networks of two types (multilabel and multiclass) are 12-channel ECG recordings with a sampling frequency of 100 and 500 Hz with normalization by the average. Machine methods work with outputs from the penultimate layers (on the best combination of all such outputs) of trained deep learning models. After training, if possible, substitutions with thresholds are selected, which allow increasing the accuracy of the model by changing the forecast at a certain threshold. Next, ensembles of all possible lengths (from 1 to the number of all models) of various types (arithmetic mean, weighted mean, mode) are constructed from all trained models, as well as artificial ones (which always return the same number, which is the number of the set of classes).

CONCLUSIONS

Training neural networks of two types (multilabel and multiclass) allows you to increase metrics.Using a combination of features from different neural networks allows you to increase the accuracy of the prediction of machine methods. The use of machine methods can either improve the metrics of the model from which features are extracted, or can be useful in ensembles. Selecting replacements at the found threshold can also improve the accuracy of methods that return a vector of probabilities. Creating ensembles from deep and machine learning methods, as well as artificial models, can improve metrics. In addition, the predictive ability of the constructed ensembles can be enhanced by the use of artificial models. The use of different types of ensembles is justified by the fact that each of them can give the highest accuracy rates in a particular set. In the future, other deep and machine learning models, types of ensembles, and tools for increasing the accuracy of models similar to finding thresholds with replacements for various medical diagnostic tasks will be considered. There are plans to develop explainable artificial intelligence.

ACKNOWLEDGEMENTS

Results obtained in numerical experiments with convolutional neural networks are supported by Ministry of Science and Education of Russian Federation, project 0729-2021-013.

RESULTS

Top 10 deep learning models that show the highest ABS:

	AlexNet	VGG16	Inception	CBB	LSTM	GRU	Xception	SCNN	LG	CNN
multilabel	67.62	67.80	66.05	**68.13**	67.25	66.65	66.65	67.16	67.39	67.53
multiclass	69.18	69.14	68.08	**70.06**	68.45	68.40	69.05	68.72	69.74	68.82

The ABS score of the best ensemble is **71.21**%.

Top 10 deep learning models (with highest ABS) after replacing with thresholds:

	AlexNet	VGG16	Inception	CBB	LSTM	GRU	Xception	SCNN	LG	CNN
multilabel	70.11	69.88	69.74	**70.34**	69.09	68.72	68.86	69.18	68.95	70.15
multiclass	70.20	69.78	68.45	**70.75**	69.28	69.00	69.69	69.46	70.15	69.32

The ABS score of the best ensemble with arficial models is **73.05**%.

Top 10 machine learning methods that show the highest ABS:

	SGD	KNN	PAC	LDA	DT	ET	ETS	RF	RC	RCC
ABS	68.77	70.55	67.99	67.34	60.66	62.00	**71.35**	70.75	68.03	68.17

The ABS score of the best ensemble with arficial models is **71.49**%.

Top 10 machine methods (with highest ABS) after replacing with thresholds:

	SGD	KNN	PAC	LDA	DT	ET	ETS	RF	RC	RCC
ABS	69.78	71.26	67.99	68.17	60.66	62.00	**71.35**	70.75	68.03	68.17

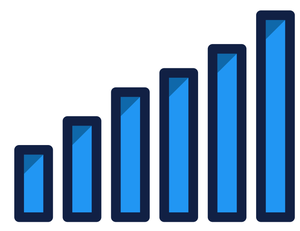
The ABS score of the best ensemble with arficial models is **72.27**%.

The best ensemble (by metric) with neural networks, machine methods and artificial models with ABS=**73.24**%.

Composition of the mode ensemble: KNN, mcCBB, mcLG, mlLSTM, mlCNN, am6, am18.

Factors that positively affect ABS:

- normalization by mean of the input data
- using a global average pooling layer
- distribution of data into samples
- using callback functions
- composition of ensembles
- using replacing with thresholds
- combining networks and machine methods
- using artificial models



DOWNLOAD