Classification of time series for small datasets

Bartłomiej Jamiołkowski Dominik Tomalczyk

January 2025

1 Introduction

The problem of time series classification concerns the analysis of signals such as data from medical devices, accelerometers or smartwatches. In particular, the analysis of this data can be used in the diagnosis of neurological and mental diseases. This project focuses on the classification of patient's daily activity for three datasets: Depresjon, Psykose and HYPERAKTIV. The data are one-dimensional time series, but the main challenge is the small number of patients in the datasets, which makes both the training and testing of models difficult.

2 Project goals

- Classification of three datasets using modern time series analysis methods.
- Evaluation of the Multiple Instance Learning (MIL) approach compared to classification of the entire data showed in this aritcle: aritcle
- Comparison of classification results of the entire data, single days and single nights.
- Benchmark of the used algorithms.

3 Datasets

We used two datasets from original article and expanded our work with third one - HYPERAKTIV.

- Psykose link
- Depresjon link
- HYPERAKTIV link

Datasets contains one-dimension time series with dates and daily brain activities.

4 Methodology

We used two feature engineering techniques proposed in article:

- M - Manual
- AM - Automated Minimal

Instead of treating data for one patient as record we did 3 splits for 24h, days(8am - 9pm) and nights(9pm - 8am). Then we treated one piece of this as a one record. For example one 24h slice is one labeled record.

We also tried to create ensemble methods:

- M ensemble - Manual (ensemble classification)
- AM ensemble - Automated Minimal (ensemble classification)

Which in practice means we take one patient, split its data into smaller pieces (24h, days or nights) make prediction for all pieces but then create a majority voting. Finally after ensemble we have one prediction per patient.

We used the same models as in original aritcle (Logistic Regression, Random Forest, SVM) and used two new (XGBoost, LGBM)

5 Results

Below we we present best accuracies from base article with model name that obtained this performance as long as our performance for ensemble method.

We cannot compare HYPERAKTIV dataset as it was not in base article. We also do not compare our method without ensemble as it is different approach and we do not have one single output for one patient

It is worth or mention that in base article they tested 5 instead of 2 approaches to feature engineering. In our approach we increased dataset approximately 20 times and added one additional dataset. Referring to that computing time has increased so we need to reduce number of tested feature engineering techniques.

Method	Clf.	Accuracy			MCC		
		24h	day	night	24h	day	night
M	LR	0.4475 ± 0.0938	0.4242 ± 0.1027	0.5337 ± 0.0281	0.0299 ± 0.0368	0.0000 ± 0.0000	0.2654 ± 0.0703
	SVM	0.6253 ± 0.0081	**0.6382 ± 0.0288**	**0.6166 ± 0.0139**	-0.0021 ± 0.0382	0.0515 ± 0.1339	**0.3430 ± 0.1159**
	RF	0.6091 ± 0.0276	0.5570 ± 0.0574	0.5336 ± 0.0333	0.1549 ± 0.0527	0.0484 ± 0.1016	0.2356 ± 0.0397
	LGBM	0.6354 ± 0.0300	0.5970 ± 0.0357	0.5000 ± 0.0300	0.1983 ± 0.0630	**0.0989 ± 0.0903**	0.2155 ± 0.0373
	XGB	**0.6657 ± 0.0059**	0.6308 ± 0.0318	0.5935 ± 0.0144	**0.2375 ± 0.0314**	0.0542 ± 0.1315	0.2620 ± 0.0710
M - ensemble	LR	0.4209 ± 0.0325	0.4685 ± 0.0607	0.6075 ± 0.0465	0.0008 ± 0.1136	0.0159 ± 0.0887	0.1903 ± 0.0884
	SVM	0.6148 ± 0.0490	0.5824 ± 0.0012	**0.6258 ± 0.0632**	0.2276 ± 0.0991	0.0000 ± 0.0000	**0.2480 ± 0.1316**
	RF	**0.7543 ± 0.0248**	0.5713 ± 0.0335	0.5535 ± 0.0261	**0.5098 ± 0.0663**	0.0641 ± 0.0757	-0.0345 ± 0.0760
	LGBM	0.7284 ± 0.0359	**0.6262 ± 0.0175**	0.4868 ± 0.0276	0.4455 ± 0.0968	**0.1982 ± 0.0451**	-0.1299 ± 0.0670
	XGB	0.6517 ± 0.0253	0.5710 ± 0.0391	0.5640 ± 0.0110	0.2825 ± 0.0967	0.0295 ± 0.1269	-0.0171 ± 0.0630
AM	LR	0.6999 ± 0.0158	0.6673 ± 0.0221	0.6923 ± 0.0391	0.4276 ± 0.0379	0.3653 ± 0.0358	0.4148 ± 0.0659
	SVM	0.7517 ± 0.0099	**0.7282 ± 0.0164**	**0.7789 ± 0.0129**	0.4348 ± 0.0182	**0.3873 ± 0.0405**	**0.5036 ± 0.0259**
	RF	0.7375 ± 0.0175	0.7195 ± 0.0205	0.7686 ± 0.0193	0.4009 ± 0.0447	0.3685 ± 0.0643	0.4788 ± 0.0417
	LGBM	**0.7592 ± 0.0152**	0.7089 ± 0.0224	0.7771 ± 0.0095	**0.4531 ± 0.0364**	0.3605 ± 0.0452	0.5030 ± 0.0221
	XGB	0.7592 ± 0.0202	0.7185 ± 0.0209	0.7751 ± 0.0095	0.4472 ± 0.0495	0.3640 ± 0.0550	0.4954 ± 0.0160
AM - ensemble	LR	0.6230 ± 0.0128	0.6842 ± 0.0727	0.6857 ± 0.0267	0.4456 ± 0.0158	0.3805 ± 0.1726	0.4272 ± 0.0223
	SVM	**0.7435 ± 0.0038**	0.7560 ± 0.0386	**0.8214 ± 0.0196**	0.4443 ± 0.0216	**0.5191 ± 0.0866**	**0.6381 ± 0.0427**
	RF	0.6913 ± 0.0277	0.7562 ± 0.0096	0.7821 ± 0.0381	**0.5633 ± 0.0452**	0.5117 ± 0.0261	0.5535 ± 0.0871
	LGBM	0.6302 ± 0.0112	**0.7598 ± 0.0151**	0.8214 ± 0.0375	0.4780 ± 0.0094	0.5057 ± 0.0321	0.6287 ± 0.0797
	XGB	0.6934 ± 0.0069	0.7346 ± 0.0322	0.7964 ± 0.0087	0.5152 ± 0.0256	0.4803 ± 0.0751	0.5913 ± 0.0205

Table 1: Table for Depresjon dataset

Method	Clf.	Accuracy			MCC		
		24h	day	night	24h	day	night
M	LR	0.5228 ± 0.0115	0.5264 ± 0.0471	0.5088 ± 0.0055	0.0474 ± 0.0222	0.0631 ± 0.1056	0.0147 ± 0.0141
	SVM	**0.5482 ± 0.0287**	**0.5595 ± 0.0371**	0.5193 ± 0.0301	**0.0970 ± 0.0572**	**0.1218 ± 0.0796**	0.0368 ± 0.0623
	RF	0.5519 ± 0.0256	0.5284 ± 0.0661	0.5193 ± 0.0465	0.1037 ± 0.0509	0.0564 ± 0.1350	0.0380 ± 0.0931
	LGBM	0.5463 ± 0.0319	0.4954 ± 0.0638	**0.5298 ± 0.0181**	0.0921 ± 0.0638	-0.0104 ± 0.1280	**0.0594 ± 0.0364**
	XGB	0.5300 ± 0.0222	0.5341 ± 0.0103	0.4965 ± 0.0446	0.0469 ± 0.0718	0.0759 ± 0.0164	-0.0133 ± 0.0902
M - ensemble	LR	0.4843 ± 0.0149	0.5389 ± 0.0195	0.5241 ± 0.0198	-0.0251 ± 0.0316	0.0805 ± 0.0419	0.0771 ± 0.0370
	SVM	**0.5594 ± 0.0344**	0.5277 ± 0.0416	**0.5736 ± 0.0183**	**0.1141 ± 0.0668**	0.0569 ± 0.0834	**0.1783 ± 0.0451**
	RF	0.5538 ± 0.0171	0.5143 ± 0.0392	0.5272 ± 0.0151	0.1048 ± 0.0331	0.0292 ± 0.0813	0.0750 ± 0.0339
	LGBM	0.4920 ± 0.0432	**0.5634 ± 0.0234**	0.5549 ± 0.0313	-0.0142 ± 0.0882	**0.1336 ± 0.0506**	0.1240 ± 0.0592
	XGB	0.5493 ± 0.0452	0.5495 ± 0.0322	0.5114 ± 0.0267	0.0952 ± 0.0897	0.1018 ± 0.0676	0.0374 ± 0.0538
AM	LR	0.5446 ± 0.0255	0.4980 ± 0.0561	**0.5664 ± 0.0372**	0.0899 ± 0.0516	0.0055 ± 0.1209	**0.1337 ± 0.0761**
	SVM	**0.5847 ± 0.0372**	0.5429 ± 0.0148	0.5428 ± 0.0320	**0.1743 ± 0.0761**	**0.0720 ± 0.0248**	0.0887 ± 0.0687
	RF	0.5046 ± 0.0415	0.5390 ± 0.0503	0.5262 ± 0.0594	0.0087 ± 0.0840	0.0657 ± 0.0986	0.0519 ± 0.1192
	LGBM	0.5191 ± 0.0395	0.5251 ± 0.0806	0.5481 ± 0.0480	0.0377 ± 0.0794	0.0417 ± 0.1663	0.0962 ± 0.0961
	XGB	0.5372 ± 0.0531	**0.5467 ± 0.0508**	0.5628 ± 0.0273	0.0785 ± 0.1126	0.0676 ± 0.1033	0.1263 ± 0.0560
AM - ensemble	LR	0.5559 ± 0.0140	0.5416 ± 0.0215	0.5485 ± 0.0109	0.1209 ± 0.0285	0.1096 ± 0.0452	0.1129 ± 0.0277
	SVM	0.6171 ± 0.0256	0.5409 ± 0.0224	0.5366 ± 0.0326	0.2341 ± 0.0539	0.1180 ± 0.0663	0.0879 ± 0.0667
	RF	**0.6211 ± 0.0304**	**0.6204 ± 0.0150**	**0.5592 ± 0.0353**	**0.2528 ± 0.0614**	**0.2660 ± 0.0324**	**0.1239 ± 0.0673**
	LGBM	0.6194 ± 0.0325	0.5939 ± 0.0228	0.5563 ± 0.0568	0.2449 ± 0.0634	0.2091 ± 0.0505	0.1191 ± 0.1111
	XGB	0.6119 ± 0.0123	0.5437 ± 0.0371	0.4994 ± 0.0255	0.2211 ± 0.0242	0.0821 ± 0.0633	0.0047 ± 0.0498

Table 2: Table for HYPERAKTIV dataset

Method	Clf.	Accuracy			MCC		
		24h	day	night	24h	day	night
M	LR	0.4305 ± 0.0581	0.3561 ± 0.0014	0.6279 ± 0.0388	0.0545 ± 0.0567	0.0000 ± 0.0000	0.2654 ± 0.0703
	SVM	**0.7009 ± 0.0367**	**0.6429 ± 0.0008**	**0.7134 ± 0.0459**	**0.2972 ± 0.0997**	-0.0109 ± 0.0219	**0.3430 ± 0.1159**
	RF	0.6012 ± 0.0251	0.5497 ± 0.0303	0.6597 ± 0.0176	0.1373 ± 0.0583	-0.0163 ± 0.0285	0.2356 ± 0.0397
	LGBM	0.6506 ± 0.0198	0.5800 ± 0.0257	0.6410 ± 0.0195	0.1972 ± 0.0433	**0.0389 ± 0.0610**	0.2155 ± 0.0373
	XGB	0.6875 ± 0.0194	**0.6342 ± 0.0176**	0.6828 ± 0.0262	0.2621 ± 0.0539	0.0016 ± 0.0622	0.2620 ± 0.0710
M - ensemble	LR	0.4222 ± 0.0319	0.4067 ± 0.0014	0.6852 ± 0.0586	-0.0284 ± 0.1119	0.0000 ± 0.0000	0.3810 ± 0.1221
	SVM	0.6704 ± 0.0181	0.5933 ± 0.0014	**0.7259 ± 0.0378**	0.3425 ± 0.0330	0.0000 ± 0.0000	**0.4413 ± 0.0867**
	RF	**0.7037 ± 0.0370**	0.5785 ± 0.0147	0.7148 ± 0.0222	**0.3827 ± 0.0929**	-0.0565 ± 0.0925	0.4317 ± 0.0662
	LGBM	0.6370 ± 0.0323	**0.6234 ± 0.0236**	0.7074 ± 0.0319	0.2298 ± 0.1218	**0.1778 ± 0.0854**	0.3882 ± 0.0790
	XGB	0.6519 ± 0.0181	0.5858 ± 0.0084	0.6481 ± 0.0000	0.2845 ± 0.0612	-0.0233 ± 0.0453	0.2748 ± 0.0222
AM	LR	0.7780 ± 0.0135	0.7405 ± 0.0349	0.7495 ± 0.0774	0.5578 ± 0.0408	0.4938 ± 0.0730	0.5586 ± 0.1088
	SVM	**0.8358 ± 0.0163**	**0.8051 ± 0.0391**	0.8312 ± 0.0403	**0.6299 ± 0.0362**	**0.5631 ± 0.0897**	0.6369 +- 0.0866
	RF	0.8160 ± 0.0133	0.7874 ± 0.0405	0.8292 ± 0.0318	0.5836 ± 0.0346	0.5263 ± 0.0871	0.6158 ± 0.0695
	LGBM	0.8197 ± 0.0207	0.7727 ± 0.0442	0.8349 ± 0.0357	0.6052 ± 0.0462	0.5090 ± 0.0885	**0.6405 ± 0.0702**
	XGB	0.8245 ± 0.0100	0.7943 ± 0.0281	**0.8397 ± 0.0310**	0.6059 ± 0.0236	0.5351 ± 0.0686	0.6452 ± 0.0625
AM - ensemble	LR	0.8395 ± 0.0104	0.6896 ± 0.0348	0.7370 ± 0.0296	0.6892 ± 0.0242	0.4370 ± 0.0528	0.5346 ± 0.0343
	SVM	0.8580 ± 0.0238	0.8428 ± 0.0388	0.8852 ± 0.0139	0.7089 ± 0.0505	0.6774 ± 0.0800	0.7626 ± 0.0296
	RF	0.8652 ± 0.0299	0.8209 ± 0.0402	0.9259 ± 0.0262	0.7276 ± 0.0618	0.6403 ± 0.0875	0.8490 ± 0.0537
	LGBM	0.8802 ± 0.0342	**0.8573 ± 0.0560**	**0.9333 ± 0.0277**	**0.7517 ± 0.0704**	**0.7095 ± 0.1178**	**0.8633 ± 0.0571**
	XGB	**0.8765 ± 0.0298**	0.8245 ± 0.0442	0.9111 ± 0.0296	0.7484 ± 0.0618	0.6445 ± 0.0983	0.8161 ± 0.0616

Table 3: Table for Psykose dataset

2*Dataset	Manual		2*Split
	Base Article	Our Ensemble	
3*Psykose	RF 0.80 ± 0.09	SVM 0.70 ± 0.03	24h
	LR 0.73 ± 0.13	LGBM 0.62 ± 0.02	Day
	RF 0.91 ± 0.12	SVM 0.72 ± 0.03	Night
3*Depresjon	LR 0.76 ± 0.16	RF 0.75 ± 0.02	24h
	LR 0.73 ± 0.13	LGBM 0.62 ± 0.01	Day
	LR 0.67 ± 0.13	SVM 0.62 ± 0.06	Night

Table 4: Performance comparison using manual feature engineering.

2*Dataset	Automated Minimal		2*Split
	Base Article	Our Ensemble	
3*Psykose	SVM 0.80 ± 0.02	XGB 0.87 ± 0.02	24h
	SVM 0.80 ± 0.06	LGBM 0.85 ± 0.05	Day
	SVM 0.91 ± 0.06	LGBM 0.93 ± 0.02	Night
3*Depresjon	RF 0.73 ± 0.06	SVM 0.74 ± 0.03	24h
	RF 0.72 ± 0.03	LGBM 0.75 ± 0.01	Day
	RF 0.74 ± 0.07	SVM 0.82 ± 0.01	Night

Table 5: Performance comparison using automated minimal feature engineering.

6 Conclusions

Experiment proved that it is worth to test ensemble approach. Using Automated Minimal feature engineering we managed to outperform results from base article for both datasets for all splits.

In terms of ensemble and splitting data for smaller parts, differences between Manual and Automated Minimal feature engineering are bigger than in data without splitting.

Data from night again seems to be more informative that data from day and from all 24 hours.

General conclusion is that using ensemble is worth testing when working with time series for small datasets as it can outperform traditional results.