

Supplementary Information

Supplementary tables and figures providing patient by patient details for each thesis chapter.

David Luke Elliott, M.Sc.



Mathematics and Statistics

Lancaster University

United Kingdom

May 2021

Chapter 3

All models performed particularly poorly on P16, P17, and P18 (see table S1). RF had particularly poor performance on P7 and P15, where no seizures were predicted despite their presence in the data, and with P16, where no seizures were present in the data but the whole record was marked as containing seizures. Indeed, all models generally performed better on records that contained seizures comparative to records without. The gains from training models on the multi-class data are mainly for pipelines ending with RF models, largely due to an increase in performance on particular patients (e.g. P16; see table S2).

Table S1: Binary performance metrics before post-processing.

Classifier	Metric	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
KNN	<i>Accuracy</i>	99.6	99.92	99.31	99.58	100	99.32	98.82	100	99.22	99.83	99.69
	<i>Sensitivity</i>	73.68	85.71	90.91	97.06	NULL	89.47	94.22	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100	100	99.53	99.7	100	99.46	99.54	100	99.22	99.83	99.69
	<i>Precision</i>	100	100	83.33	94.29	NULL	70.83	97.02	NULL	NULL	NULL	NULL
	<i>F1-score</i>	84.85	92.31	86.96	95.65	NULL	79.07	95.6	NULL	NULL	NULL	NULL
	<i>AUC</i>	92.07	99.97	95.16	99.1	NULL	96.92	97.44	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0	0	16.49	10.18	0	19.02	14.2	0	28.23	5.99	11.2
<i>CPU Time</i>	1.55	1.69	1.97	1.73	1.47	1.52	1.25	1.83	1.69	2.47	2.83	
RF	<i>Accuracy</i>	99.92	99.92	98.93	99.65	100	98.19	86.36	99.1	98.86	99.5	95.72
	<i>Sensitivity</i>	94.74	100	90.91	97.06	NULL	94.74	0	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100	99.91	99.14	99.78	100	98.24	100	99.1	98.86	99.5	95.72
	<i>Precision</i>	100	87.5	73.17	95.65	NULL	43.9	0	NULL	NULL	NULL	NULL
	<i>F1-score</i>	97.3	93.33	81.08	96.35	NULL	60	0	NULL	NULL	NULL	NULL
	<i>AUC</i>	100	100	95.91	99.51	NULL	98.93	50	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0	3.05	30.23	7.63	0	62.49	0	32.54	41.05	17.97	153.97
<i>CPU Time</i>	0.44	0.19	0.78	0.06	0.5	0.05	0.81	0.84	0.8	0.31	0.56	
SVM	<i>Accuracy</i>	99.92	99.75	98.55	99.51	100	99.4	98.74	99.84	99.14	99.75	97.43
	<i>Sensitivity</i>	94.74	100	90.91	92.65	NULL	89.47	96.53	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100	99.74	98.75	99.85	100	99.54	99.09	99.84	99.14	99.75	97.43
	<i>Precision</i>	100	70	65.22	96.92	NULL	73.91	94.35	NULL	NULL	NULL	NULL
	<i>F1-score</i>	97.3	82.35	75.95	94.74	NULL	80.95	95.43	NULL	NULL	NULL	NULL
	<i>AUC</i>	100	100	96.35	99.97	NULL	99.08	98.5	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0	9.16	43.97	5.09	0	16.3	28.39	5.92	30.79	8.99	92.38
<i>CPU Time</i>	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.22	0.12	
SMV	<i>Accuracy</i>	99.92	100	98.93	99.65	100	99.25	98.9	100	98.93	99.67	99.07
	<i>Sensitivity</i>	94.74	100	90.91	97.06	NULL	94.74	94.8	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100	100	99.14	99.78	100	99.31	99.54	100	98.93	99.67	99.07
	<i>Precision</i>	100	100	73.17	95.65	NULL	66.67	97.04	NULL	NULL	NULL	NULL
	<i>F1-score</i>	97.3	100	81.08	96.35	NULL	78.26	95.91	NULL	NULL	NULL	NULL
	<i>AUC</i>	100	100	95.66	99.96	NULL	99.52	98.58	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0	0	30.23	7.63	0	24.45	14.2	0	38.49	11.98	33.59
<i>CPU Time</i>	2.19	1.95	2.59	1.77	2.19	1.78	2.05	2.59	2.8	2.69	3.47	
Classifier	Metric	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	Ave
KNN	<i>Accuracy</i>	98.71	99.18	99.31	99.83	95.76	95.99	90.25	99.71	99.7	99.66	98.73
	<i>Sensitivity</i>	100	NULL	80.26	89.47	NULL	100	NULL	94.12	94.44	100	91.49
	<i>Specificity</i>	98.67	99.18	99.8	100	95.76	95.75	90.25	99.78	99.92	99.66	98.85
	<i>Precision</i>	70.83	NULL	91.04	100	NULL	58.65	NULL	84.21	98.08	58.33	85.13
	<i>F1-score</i>	82.93	NULL	85.31	94.44	NULL	73.94	NULL	88.89	96.23	73.68	86.91
	<i>AUC</i>	99.34	NULL	91.98	94.71	NULL	97.88	NULL	96.11	97.17	99.83	96.74
	<i>FP/h</i>	46.28	29.42	7.06	0	152.82	144.27	351.03	7.93	2.73	12.2	40.91
<i>CPU Time</i>	1.33	1.67	2.83	1.39	2.58	1.2	2.7	1.64	1.62	1.59	1.84	
RF	<i>Accuracy</i>	95.5	99.03	99.25	98.43	0	92.82	88.33	97.87	99.17	98.71	92.63
	<i>Sensitivity</i>	100	NULL	84.21	0	NULL	98.36	NULL	94.12	98.15	100	80.94
	<i>Specificity</i>	95.36	99.03	99.63	100	0	92.49	88.33	97.92	99.21	98.71	93.38
	<i>Precision</i>	40.96	NULL	85.33	0	NULL	44.12	NULL	36.36	84.13	26.92	55.23
	<i>F1-score</i>	58.12	NULL	84.77	0	NULL	60.91	NULL	52.46	90.6	42.42	62.87
	<i>AUC</i>	98.71	NULL	94.71	50	NULL	99.35	NULL	99.41	99.83	99.98	91.26
	<i>FP/h</i>	161.98	34.77	12.94	0	3600	254.99	420.21	74.01	27.27	46.34	237.21
<i>CPU Time</i>	0.56	0.84	0.64	0.5	0.62	0.06	0.22	0.23	0.17	0.38	0.46	
SVM	<i>Accuracy</i>	98.35	98.81	99.25	99.75	86.78	94.04	94.59	99.63	99.62	93.9	97.94
	<i>Sensitivity</i>	100	NULL	84.21	89.47	NULL	88.52	NULL	70.59	90.74	100	91.37
	<i>Specificity</i>	98.29	98.81	99.63	99.92	86.78	94.37	94.59	100	100	93.87	98.07
	<i>Precision</i>	65.38	NULL	85.33	94.44	NULL	48.65	NULL	100	100	7.22	77.03
	<i>F1-score</i>	79.07	NULL	84.77	91.89	NULL	62.79	NULL	82.76	95.15	13.46	79.74
	<i>AUC</i>	99.81	NULL	98.01	99.79	NULL	98.75	NULL	99.66	99.96	100	99.22
	<i>FP/h</i>	59.5	42.79	12.94	2.98	476.08	191.24	194.73	0	0	219.51	68.61
<i>CPU Time</i>	0.08	0.05	0.23	0.06	0.06	0.05	0.06	0.05	0.06	0.08	0.08	
SMV	<i>Accuracy</i>	98.35	98.81	99.28	99.83	90.69	92.73	88.97	99.78	99.47	98.64	98.14
	<i>Sensitivity</i>	100	NULL	84.21	89.47	NULL	100	NULL	94.12	94.44	100	94.96
	<i>Specificity</i>	98.29	98.81	99.66	100	90.69	92.29	88.97	99.85	99.68	98.64	98.21
	<i>Precision</i>	65.38	NULL	86.49	100	NULL	43.88	NULL	88.89	92.73	25.93	79.68
	<i>F1-score</i>	79.07	NULL	85.33	94.44	NULL	61	NULL	91.43	93.58	41.18	84.23
	<i>AUC</i>	99.81	NULL	95.78	99.74	NULL	99.76	NULL	98.07	99.92	99.99	98.98
	<i>FP/h</i>	59.5	42.79	11.76	0	335.02	261.7	397.15	5.29	10.91	48.78	63.5
<i>CPU Time</i>	2.14	2.44	3.2	2.06	3.2	1.44	2.97	1.92	1.73	2.27	2.35	

Note. Metrics are in percent, apart from FP/h and CPU Time. The best average score for each metric, across classifiers, are in bold. NULL is where a metric cannot be calculated due to there being no seizures present in the data.

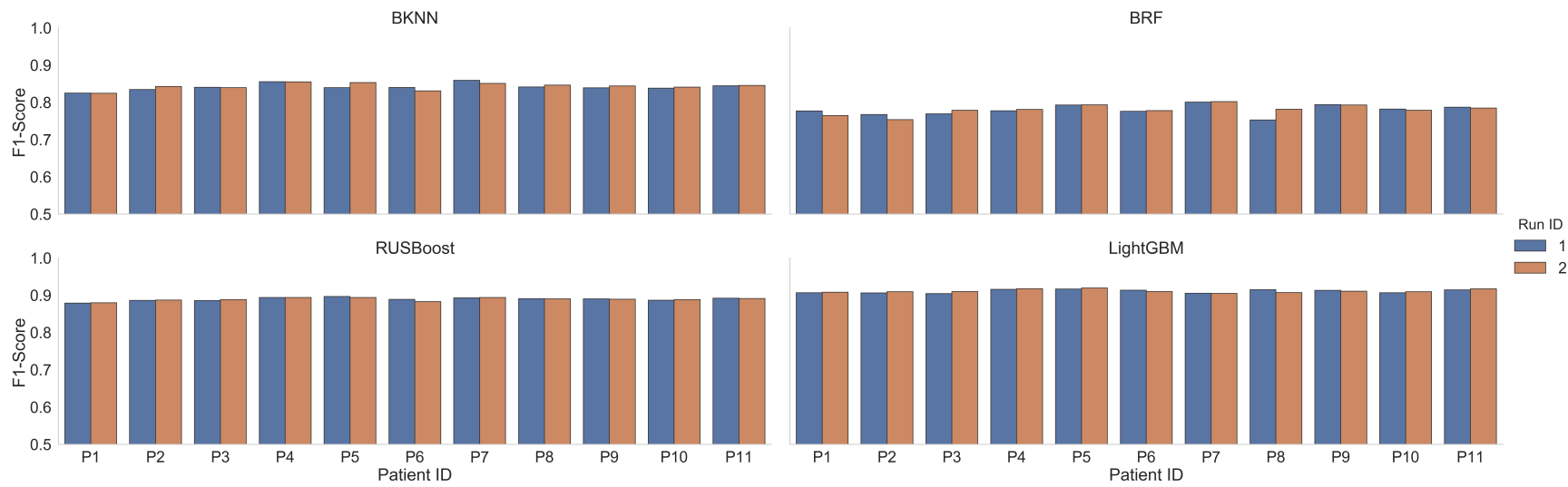
Table S2: Multiclass performance metrics before post-processing.

Classifier	Metric	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
KNN	<i>Accuracy</i>	99.92	100	99.24	99.36	100	99.55	98.9	99.84	99.22	99.75	96.19
	<i>Sensitivity</i>	94.74	100	93.94	95.77	NULL	95.24	95.43	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100	100	99.37	99.55	100	99.62	99.45	99.84	99.22	99.75	96.19
	<i>Precision</i>	100	100	79.49	91.89	NULL	80	96.53	NULL	NULL	NULL	NULL
	<i>F1-score</i>	97.3	100	86.11	93.79	NULL	86.96	95.98	NULL	NULL	NULL	NULL
	<i>AUC</i>	99.62	100	96.4	98.35	NULL	97.02	98.22	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0	0	21.98	15.27	0	13.58	17.03	5.92	28.23	8.99	137.17
	<i>CPU Time</i>	1.66	2.84	2.77	2.28	2.28	2.31	2.2	1.52	1.41	3.28	1.53
RF	<i>Accuracy</i>	99.84	99.92	98.7	99.43	100	98.64	98.74	99.59	98.5	99.58	97.59
	<i>Sensitivity</i>	89.47	100	90.91	95.77	NULL	95.24	95.43	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100	99.91	98.9	99.63	100	98.7	99.27	99.59	98.5	99.58	97.59
	<i>Precision</i>	100	87.5	68.18	93.15	NULL	54.05	95.43	NULL	NULL	NULL	NULL
	<i>F1-score</i>	94.44	93.33	77.92	94.44	NULL	68.97	95.43	NULL	NULL	NULL	NULL
	<i>AUC</i>	100	100	96.06	99.88	NULL	99.34	98.66	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0	3.05	38.47	12.72	0	46.19	22.71	14.79	53.88	14.98	86.78
	<i>CPU Time</i>	0.7	0.45	0.28	0.39	0.28	0.25	0.23	0.16	0.39	1.03	0.19
SVM	<i>Accuracy</i>	99.84	99.75	98.63	99.29	99.92	99.4	98.74	99.26	95.94	98.67	96.35
	<i>Sensitivity</i>	89.47	100	93.94	94.37	NULL	95.24	96	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100	99.74	98.75	99.55	99.92	99.46	99.18	99.26	95.94	98.67	96.35
	<i>Precision</i>	100	70	65.96	91.78	NULL	74.07	94.92	NULL	NULL	NULL	NULL
	<i>F1-score</i>	94.44	82.35	77.5	93.06	NULL	83.33	95.45	NULL	NULL	NULL	NULL
	<i>AUC</i>	99.79	99.9	95.96	99.56	NULL	99.53	97.97	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0	9.16	43.97	15.27	2.91	19.02	25.55	26.62	146.26	47.92	131.57
	<i>CPU Time</i>	0.3	0.92	0.23	0.91	0.25	0.27	0.22	0.27	0.31	0.19	0.28
SMV	<i>Accuracy</i>	99.92	99.75	98.63	99.43	100	99.32	98.74	99.1	97.65	99.92	95.65
	<i>Sensitivity</i>	94.74	100	93.94	97.18	NULL	95.24	95.43	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100	99.74	98.75	99.55	100	99.39	99.27	99.1	97.65	99.92	95.65
	<i>Precision</i>	100	70	65.96	92	NULL	71.43	95.43	NULL	NULL	NULL	NULL
	<i>F1-score</i>	97.3	82.35	77.5	94.52	NULL	81.63	95.43	NULL	NULL	NULL	NULL
	<i>AUC</i>	100	100	95.74	99.51	NULL	99.39	98.52	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0	9.16	43.97	15.27	0	21.74	22.71	32.54	84.68	3	156.77
	<i>CPU Time</i>	2.62	4.27	3.42	3.61	2.83	2.77	2.55	1.95	2.11	4.47	1.97
Classifier	Metric	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	Ave
KNN	<i>Accuracy</i>	95.5	99.11	99.28	99.75	94.04	95.15	89.18	99.71	99.7	99.39	98.23
	<i>Sensitivity</i>	97.22	NULL	82.89	89.47	NULL	95.31	NULL	77.78	94.55	87.5	92.3
	<i>Specificity</i>	95.44	99.11	99.7	99.92	94.04	95.14	89.18	100	99.92	99.46	98.33
	<i>Precision</i>	42.17	NULL	87.5	94.44	NULL	55.45	NULL	100	98.11	46.67	82.48
	<i>F1-score</i>	58.82	NULL	85.14	91.89	NULL	70.11	NULL	87.5	96.3	60.87	85.44
	<i>AUC</i>	98.92	NULL	93.22	97.27	NULL	98.19	NULL	99.18	98.02	98.8	97.94
	<i>FP/h</i>	158.68	32.1	10.58	2.98	214.53	164.4	389.47	0	2.73	19.51	59.2
	<i>CPU Time</i>	1.59	3.08	4.44	1.22	3.12	0.98	1.52	3.02	3.05	2.25	2.3
RF	<i>Accuracy</i>	95.78	99.18	99.25	99.17	95.92	95.99	89.04	99.34	99.7	98.98	98.23
	<i>Sensitivity</i>	97.22	NULL	84.21	94.74	NULL	100	NULL	88.89	98.18	87.5	93.66
	<i>Specificity</i>	95.73	99.18	99.63	99.24	95.92	95.74	89.04	99.48	99.76	99.05	98.31
	<i>Precision</i>	43.75	NULL	85.33	66.67	NULL	59.81	NULL	69.57	94.74	33.33	73.19
	<i>F1-score</i>	60.34	NULL	84.77	78.26	NULL	74.85	NULL	78.05	96.43	48.28	80.42
	<i>AUC</i>	99.32	NULL	96.36	99.75	NULL	99.85	NULL	96.14	99.97	99.73	98.85
	<i>FP/h</i>	148.76	29.42	12.94	26.84	146.94	144.27	394.59	18.5	8.18	34.15	59.91
	<i>CPU Time</i>	0.31	0.48	0.67	0.36	0.27	0.2	0.52	0.34	0.25	0.42	0.39
SVM	<i>Accuracy</i>	94.58	99.26	99.22	99.92	99.1	96.37	96.94	99.56	99.62	98.58	98.52
	<i>Sensitivity</i>	61.11	NULL	84.21	94.74	NULL	43.75	NULL	77.78	96.36	100	86.69
	<i>Specificity</i>	95.73	99.26	99.6	100	99.1	99.7	96.94	99.85	99.76	98.57	98.82
	<i>Precision</i>	32.84	NULL	84.21	100	NULL	90.32	NULL	87.5	94.64	27.59	77.99
	<i>F1-score</i>	42.72	NULL	84.21	97.3	NULL	58.95	NULL	82.35	95.5	43.24	79.26
	<i>AUC</i>	95.31	NULL	96.66	98.54	NULL	94.19	NULL	97.64	99.95	99.82	98.06
	<i>FP/h</i>	148.76	26.75	14.11	0	32.33	10.07	110.18	5.29	8.18	51.22	41.67
	<i>CPU Time</i>	0.23	0.94	1.16	0.23	0.91	0.22	0.92	0.3	0.97	0.91	0.52
SMV	<i>Accuracy</i>	94.86	99.26	99.25	99.83	95.02	96.27	88.9	99.71	99.77	98.58	98.07
	<i>Sensitivity</i>	97.22	NULL	85.53	94.74	NULL	96.88	NULL	88.89	96.36	87.5	94.13
	<i>Specificity</i>	94.78	99.26	99.6	99.92	95.02	96.23	88.9	99.85	99.92	98.64	98.15
	<i>Precision</i>	38.89	NULL	84.42	94.74	NULL	62	NULL	88.89	98.15	25.93	75.99
	<i>F1-score</i>	55.56	NULL	84.97	94.74	NULL	75.61	NULL	88.89	97.25	40	81.98
	<i>AUC</i>	98.92	NULL	96.4	99.53	NULL	99.65	NULL	98.6	99.93	99.8	98.92
	<i>FP/h</i>	181.82	26.75	14.11	2.98	179.27	127.49	399.72	5.29	2.73	48.78	65.65
	<i>CPU Time</i>	2.17	4.36	6.11	2	4.36	1.47	2.94	3.62	4.42	3.7	3.22

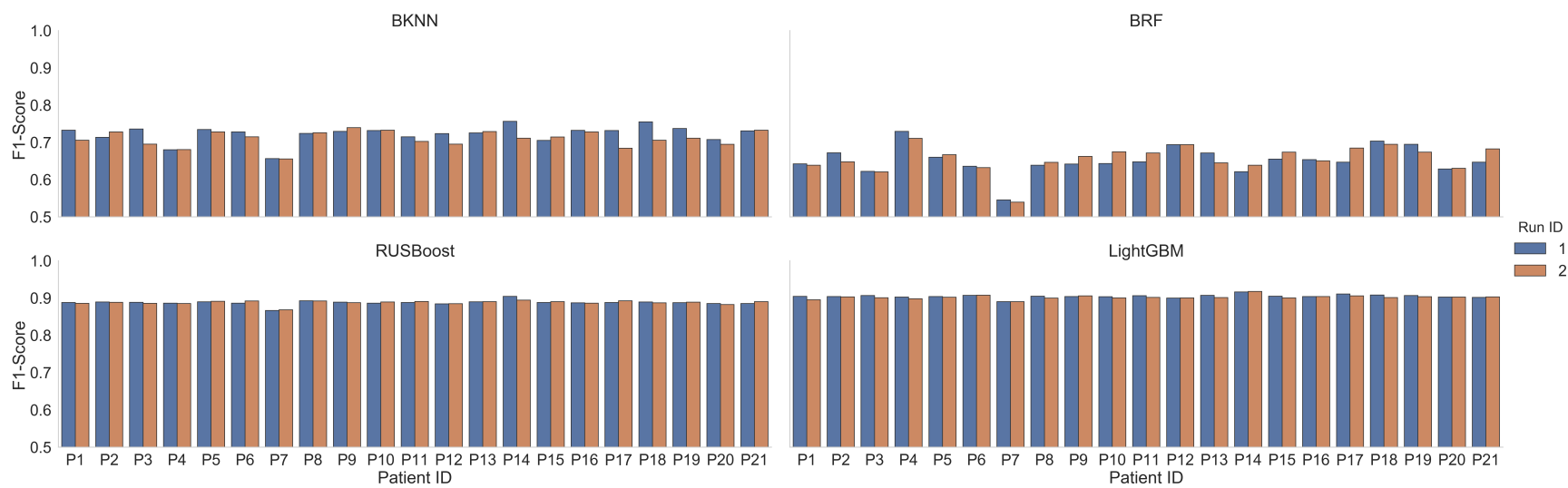
Note. Metrics are in percent, apart from FP/h and CPU Time. The best average score for each metric, across classifiers, are in bold. NULL is where a metric cannot be calculated due to there being no seizures present in the data.

Chapter 4

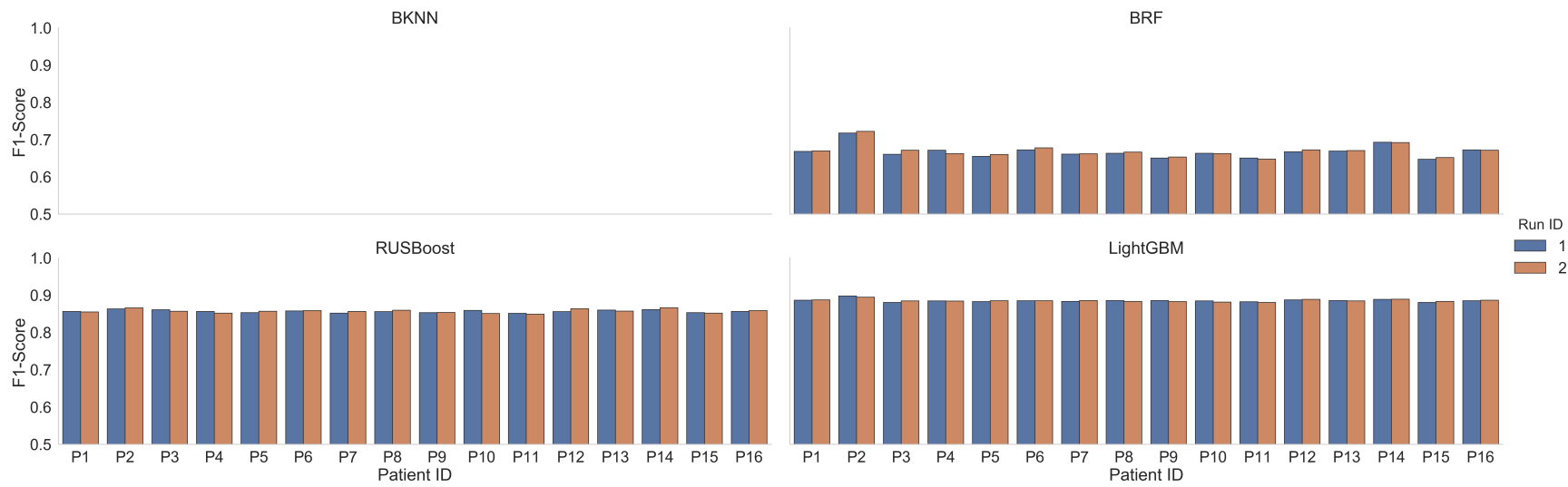
For the TUH (Absence) dataset, there was particularly poor performance on P7 (see figure S3), due to marking seizures at the beginning of the record where there is a prolonged period of relative inactivity, and not fully marking seizures where present. For the NHS (Preston) dataset, P18 generally had a lot of false positives, similar to classical models in chapter 3, both when trained alone and when combined with the NHS (Leeds) dataset. This is due to the large electrical artefacts present in this recording which are misclassified. There was also poor sensitivity on P14, P17, P19 in the NHS (Preston) dataset, but these were improved when trained in combination with the NHS (Leeds) dataset (see figure 4.A.4). Additionally, combining NHS (Preston) and NHS (Leeds) datasets improved the false positive rate for P2 in the NHS (Leeds) dataset, but did not improve the poor sensitivity for P14 (see tables S3 - S6).



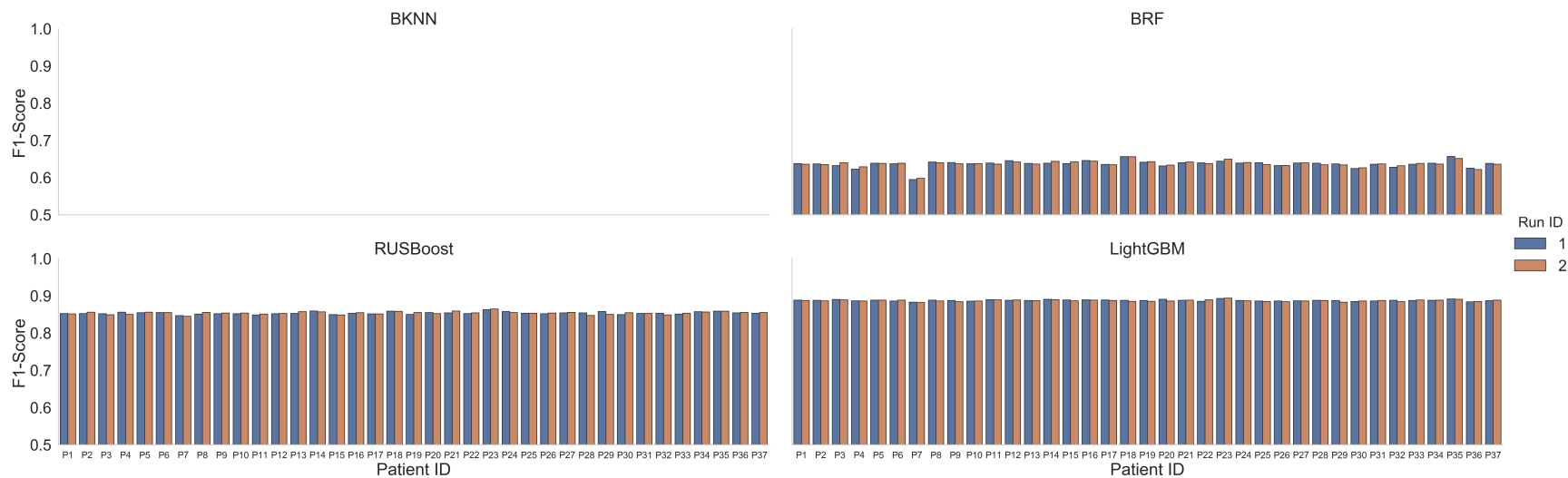
(a) TUH



(b) Preston

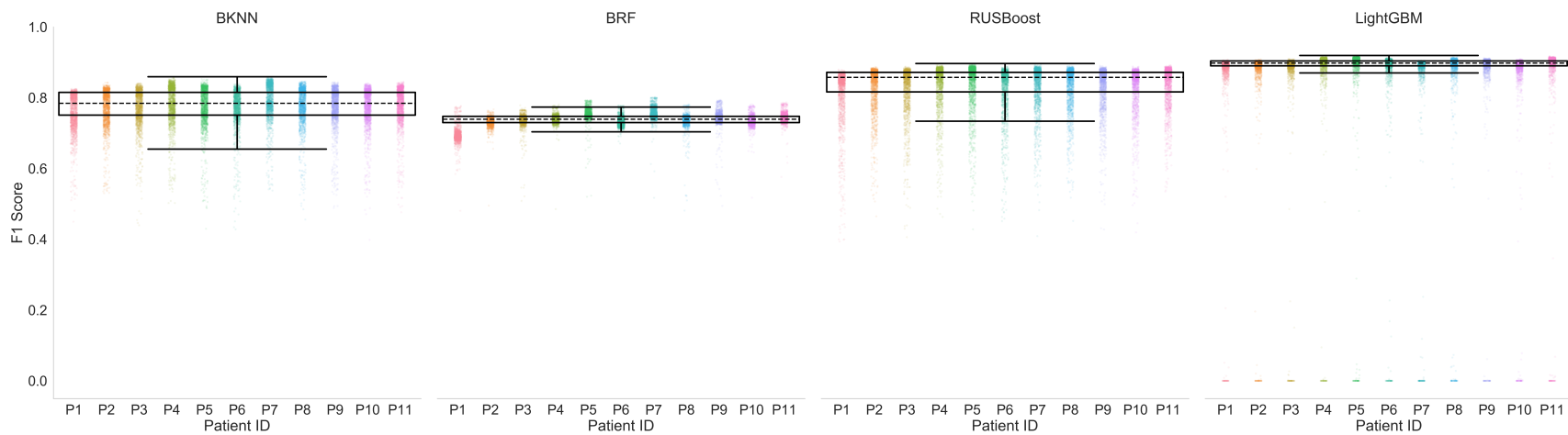


(c) Leeds

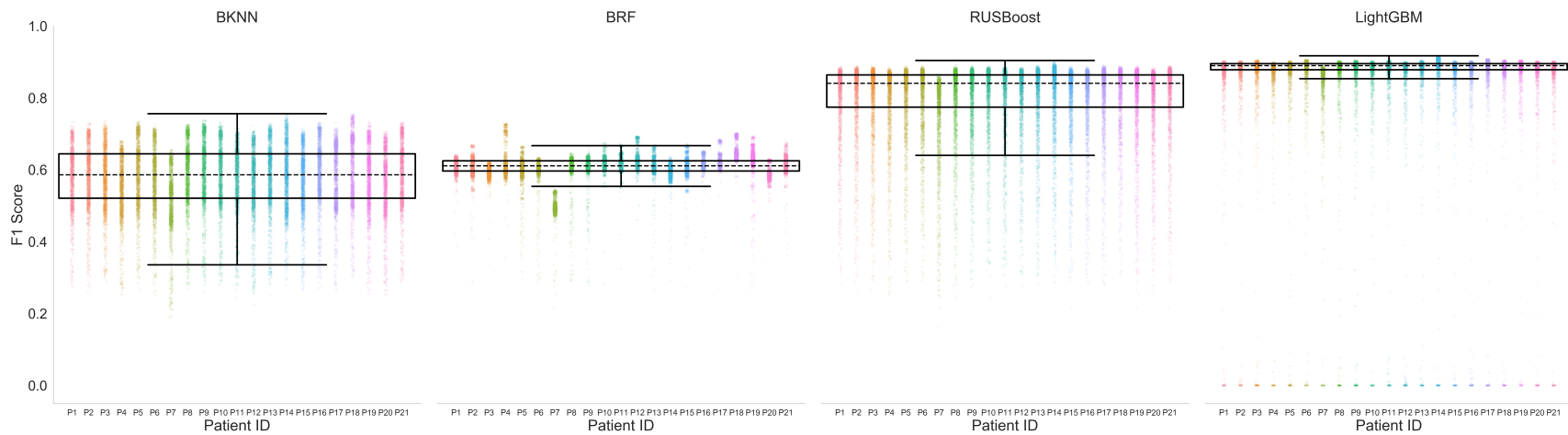


(d) Combined

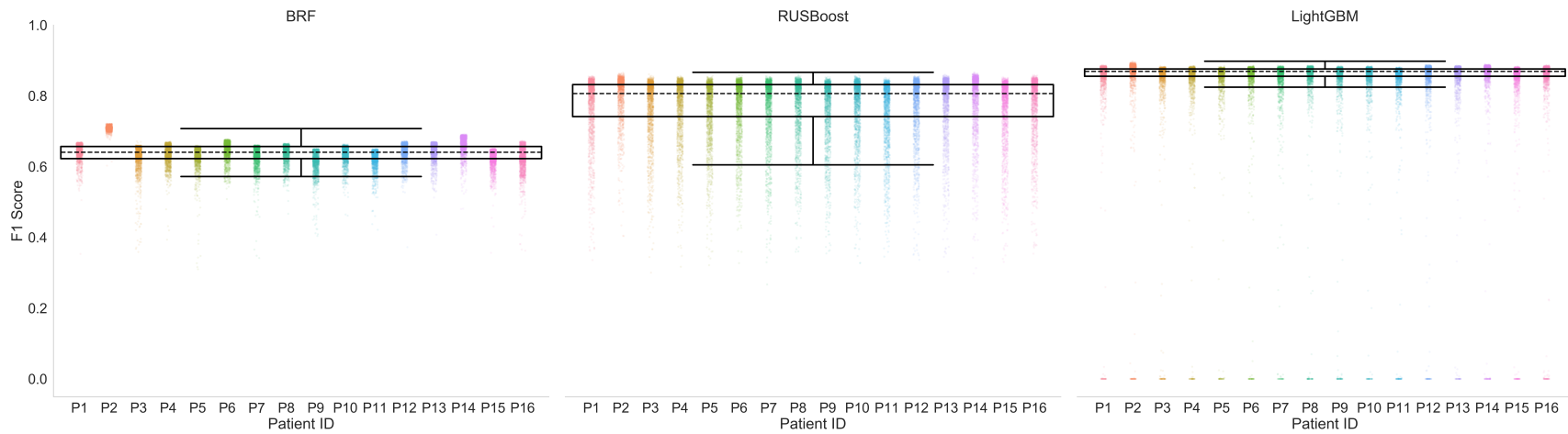
Figure S1: Best validation scores for each random state during Bayesian optimisation across patients and classifiers.



(a) TUH

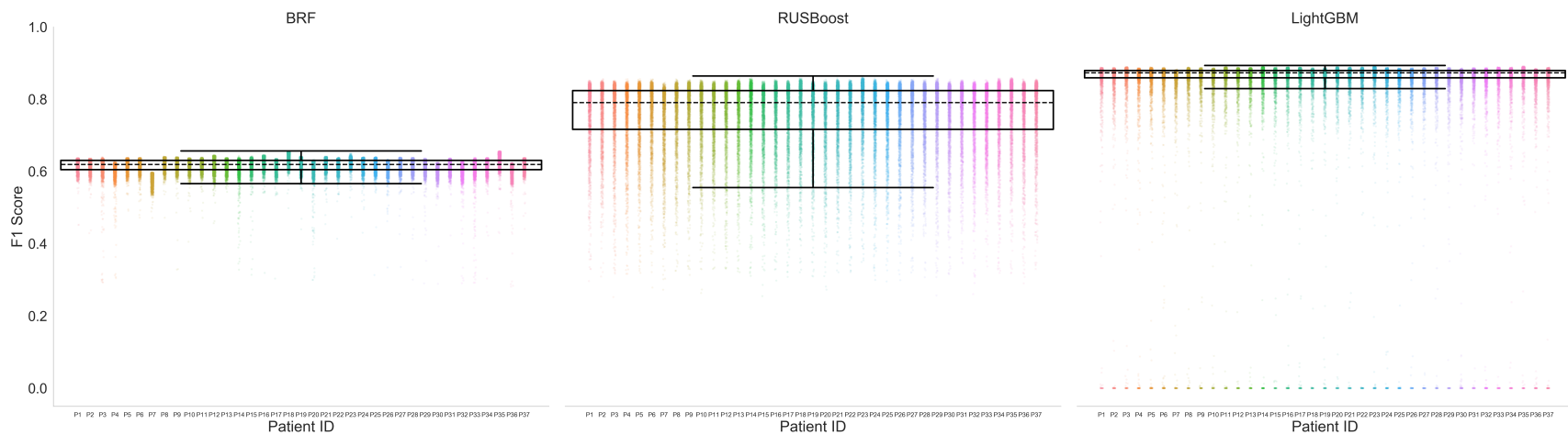


(b) Preston



(c) Leads

∞



(d) Combined

Figure S2: All Bayesian optimisation validation F1-scores across patients and classifiers.

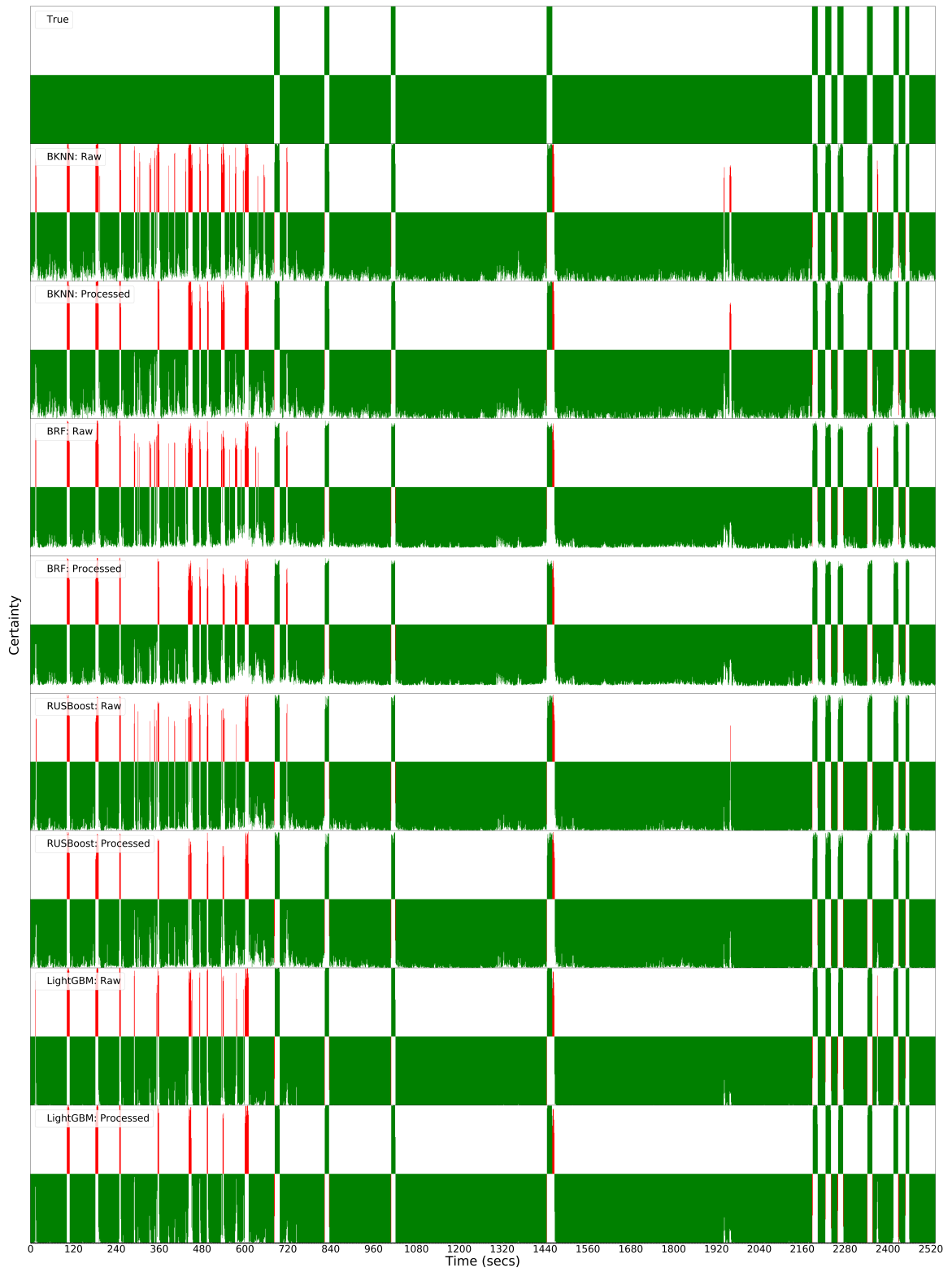


Figure S3: Prediction certainty of models on a patient record (P7), from the TUH (Absence) dataset, where performance was poor.

Note. All models mark seizures at the beginning of the record where there is a prolonged period of relative inactivity. Furthermore, all models mark the seizure in the middle of the record for longer than present and the boosted models do not fully mark the seizures towards the end of the record.

Table S3: TUH (Absence) performance metrics before post-processing.

Classifier	Metric	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
BKNN	<i>Accuracy</i>	97.69	98.68	99.25	98.23	95.86	98.09	94.75	99.37	98.53	99.75	98.4
	<i>Sensitivity</i>	82.33	88.99	89.77	73.81	74.51	84.92	91.39	85.71	86.96	94.59	78.79
	<i>Specificity</i>	99.48	99.09	99.52	99.35	97.75	98.93	94.96	99.46	98.74	99.91	98.88
	<i>Precision</i>	94.91	80.17	84.04	83.78	74.51	83.59	53.49	50.0	55.56	97.22	63.41
	<i>F1-score</i>	88.17	84.35	86.81	78.48	74.51	84.25	67.48	63.16	67.8	95.89	70.27
	<i>AUC</i>	97.06	95.46	99.38	91.79	90.55	94.29	96.96	99.83	96.76	98.07	93.67
	<i>FP/h</i>	16.66	31.6	16.78	22.5	74.46	36.15	170.55	19.3	44.51	2.99	39.3
	<i>CPU Time</i>	7.91	7.69	7.92	2.89	0.91	9.16	11.53	2.94	3.55	3.81	5.06
BRF	<i>Accuracy</i>	97.94	96.85	99.44	97.92	95.86	96.84	95.03	99.11	94.36	99.25	97.74
	<i>Sensitivity</i>	83.94	93.58	88.64	73.81	85.29	84.13	90.07	100.0	86.96	91.89	78.79
	<i>Specificity</i>	99.58	96.99	99.74	99.02	96.8	97.66	95.34	99.1	94.49	99.49	98.21
	<i>Precision</i>	95.87	56.35	90.7	77.5	70.16	69.74	55.06	41.18	22.22	85.0	52.0
	<i>F1-score</i>	89.51	70.34	89.66	75.61	76.99	76.26	68.34	58.33	35.4	88.31	62.65
	<i>AUC</i>	97.23	98.58	98.26	93.53	97.2	97.49	97.5	99.9	97.85	98.95	95.64
	<i>FP/h</i>	13.63	104.02	8.95	33.75	105.97	79.2	157.76	32.17	194.74	17.91	62.88
	<i>CPU Time</i>	0.89	0.88	0.05	0.48	0.89	0.36	0.3	0.58	0.09	0.62	1.05
RUSBoost	<i>Accuracy</i>	98.02	98.72	99.57	98.54	96.66	98.61	96.17	98.84	99.3	99.5	98.84
	<i>Sensitivity</i>	81.93	90.83	85.23	69.05	66.67	79.37	89.4	85.71	82.61	94.59	72.73
	<i>Specificity</i>	99.91	99.05	99.97	99.89	99.31	99.85	96.6	98.92	99.61	99.66	99.48
	<i>Precision</i>	99.03	79.84	98.68	96.67	89.47	97.09	62.5	33.33	79.17	89.74	77.42
	<i>F1-score</i>	89.67	84.98	91.46	80.56	76.4	87.34	73.57	48.0	80.85	92.11	75.0
	<i>AUC</i>	98.42	97.77	98.38	92.13	96.69	97.89	96.47	99.09	99.71	97.62	93.43
	<i>FP/h</i>	3.03	32.92	1.12	3.75	22.91	5.16	115.12	38.61	13.91	11.94	18.34
	<i>CPU Time</i>	1.83	1.77	2.31	0.81	1.0	1.42	2.05	0.94	1.03	1.19	1.11
LightGBM	<i>Accuracy</i>	98.19	98.87	99.69	98.23	97.3	98.66	96.57	99.82	99.23	99.59	98.76
	<i>Sensitivity</i>	83.94	88.99	89.77	73.81	68.63	78.57	88.08	71.43	86.96	91.89	72.73
	<i>Specificity</i>	99.86	99.28	99.97	99.35	99.83	99.95	97.1	100.0	99.45	99.83	99.4
	<i>Precision</i>	98.58	83.62	98.75	83.78	97.22	99.0	65.84	100.0	74.07	94.44	75.0
	<i>F1-score</i>	90.67	86.22	94.05	78.48	80.46	87.61	75.35	83.33	80.0	93.15	73.85
	<i>AUC</i>	97.45	98.83	99.07	91.6	96.87	98.95	98.31	99.74	99.8	98.49	96.28
	<i>FP/h</i>	4.54	25.02	1.12	22.5	5.73	1.72	98.07	0.0	19.47	5.97	20.96
	<i>CPU Time</i>	0.09	0.25	0.22	0.09	0.09	0.08	0.02	0.09	0.11	0.14	0.08

Note. All metrics are in percent apart from FP/h and CPU Time.

Table S4: NHS (Preston) performance metrics before post-processing.

Classifier	Metric	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
BKNN	<i>Accuracy</i>	98.89	99.15	98.32	98.87	99.76	98.79	96.77	98.19	96.08	99.67	96.81
	<i>Sensitivity</i>	85.71	88.89	82.93	89.74	NULL	75.86	78.84	NULL	NULL	NULL	NULL
	<i>Specificity</i>	99.12	99.23	98.82	99.4	99.76	99.31	99.91	98.19	96.08	99.67	96.81
	<i>Precision</i>	62.07	47.06	69.39	89.74	NULL	70.97	99.33	NULL	NULL	NULL	NULL
	<i>F1-score</i>	72.0	61.54	75.56	89.74	NULL	73.33	87.91	NULL	NULL	NULL	NULL
	<i>AUC</i>	96.34	99.15	92.36	97.25	NULL	90.34	95.15	NULL	NULL	NULL	NULL
	<i>FP/h</i>	31.3	27.48	41.22	20.35	8.74	24.45	2.84	65.08	141.13	11.98	114.77
	<i>CPU Time</i>	2.56	2.22	2.05	3.05	2.36	3.19	1.44	1.45	3.3	1.86	2.5
BRF	<i>Accuracy</i>	98.66	98.64	98.32	98.3	99.43	98.57	97.95	95.32	95.44	98.75	92.92
	<i>Sensitivity</i>	85.71	88.89	87.8	76.92	NULL	79.31	86.77	NULL	NULL	NULL	NULL
	<i>Specificity</i>	98.87	98.72	98.66	99.55	99.43	99.0	99.91	95.32	95.44	98.75	92.92
	<i>Precision</i>	56.25	34.78	67.92	90.91	NULL	63.89	99.39	NULL	NULL	NULL	NULL
	<i>F1-score</i>	67.92	50.0	76.6	83.33	NULL	70.77	92.66	NULL	NULL	NULL	NULL
	<i>AUC</i>	97.69	95.65	95.58	98.35	NULL	95.21	97.18	NULL	NULL	NULL	NULL
	<i>FP/h</i>	39.84	45.8	46.72	15.27	20.4	35.32	2.84	168.61	164.22	44.93	254.74
	<i>CPU Time</i>	0.95	0.45	1.06	0.61	1.28	0.09	0.19	0.02	0.0	0.69	0.5
RUSBoost	<i>Accuracy</i>	99.37	99.92	99.01	99.22	100.0	99.4	96.45	100.0	100.0	99.83	99.77
	<i>Sensitivity</i>	66.67	88.89	85.37	88.46	NULL	75.86	76.19	NULL	NULL	NULL	NULL
	<i>Specificity</i>	99.92	100.0	99.45	99.85	100.0	99.92	100.0	100.0	100.0	99.83	99.77
	<i>Precision</i>	93.33	100.0	83.33	97.18	NULL	95.65	100.0	NULL	NULL	NULL	NULL
	<i>F1-score</i>	77.78	94.12	84.34	92.62	NULL	84.62	86.49	NULL	NULL	NULL	NULL
	<i>AUC</i>	96.25	97.75	96.61	97.69	NULL	92.03	96.02	NULL	NULL	NULL	NULL
	<i>FP/h</i>	2.85	0.0	19.24	5.09	0.0	2.72	0.0	0.0	0.0	5.99	8.4
	<i>CPU Time</i>	0.64	0.94	0.92	1.05	0.78	0.7	0.89	0.75	0.84	0.78	0.88
LightGBM	<i>Accuracy</i>	99.68	99.83	99.31	98.02	100.0	99.32	97.4	100.0	100.0	99.83	99.84
	<i>Sensitivity</i>	80.95	77.78	85.37	64.1	NULL	75.86	82.54	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100.0	100.0	99.76	100.0	100.0	99.85	100.0	100.0	100.0	99.83	99.84
	<i>Precision</i>	100.0	100.0	92.11	100.0	NULL	91.67	100.0	NULL	NULL	NULL	NULL
	<i>F1-score</i>	89.47	87.5	88.61	78.13	NULL	83.02	90.43	NULL	NULL	NULL	NULL
	<i>AUC</i>	97.38	95.09	94.52	96.88	NULL	96.04	96.18	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0.0	0.0	8.24	0.0	0.0	5.43	0.0	0.0	0.0	5.99	5.6
	<i>CPU Time</i>	0.02	0.12	0.06	0.12	0.0	0.08	0.08	0.02	0.11	0.11	0.12
Classifier	Metric	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	
BKNN	<i>Accuracy</i>	97.52	97.77	98.43	98.76	94.94	93.66	90.53	99.78	98.86	98.31	
	<i>Sensitivity</i>	92.11	NULL	70.37	95.24	NULL	30.43	NULL	89.47	83.33	100.0	
	<i>Specificity</i>	97.72	97.77	99.46	98.82	94.94	98.01	90.53	99.93	99.6	98.3	
	<i>Precision</i>	59.32	NULL	82.61	58.82	NULL	51.22	NULL	94.44	90.91	26.47	
	<i>F1-score</i>	72.16	NULL	76.0	72.73	NULL	38.18	NULL	91.89	86.96	41.86	
	<i>AUC</i>	99.13	NULL	87.27	99.42	NULL	92.56	NULL	97.16	94.68	99.79	
	<i>FP/h</i>	79.34	80.24	18.82	41.76	182.2	67.1	340.78	2.64	13.64	60.98	
	<i>CPU Time</i>	1.44	2.75	2.66	1.66	2.62	1.2	2.84	2.14	1.73	3.05	
BRF	<i>Accuracy</i>	97.06	97.62	98.33	99.25	95.18	95.71	88.26	97.87	99.47	96.95	
	<i>Sensitivity</i>	57.89	NULL	79.63	95.24	NULL	49.28	NULL	21.05	88.33	100.0	
	<i>Specificity</i>	98.48	97.62	99.02	99.33	95.18	98.9	88.26	98.96	100.0	96.93	
	<i>Precision</i>	57.89	NULL	74.78	71.43	NULL	75.56	NULL	22.22	100.0	16.67	
	<i>F1-score</i>	57.89	NULL	77.13	81.63	NULL	59.65	NULL	21.62	93.81	28.57	
	<i>AUC</i>	98.03	NULL	90.94	99.16	NULL	95.71	NULL	97.59	96.78	99.51	
	<i>FP/h</i>	52.89	85.59	34.11	23.86	173.39	36.91	422.78	37.0	0.0	109.76	
	<i>CPU Time</i>	0.52	0.48	0.38	0.3	0.34	0.28	0.34	0.53	0.72	0.62	
RUSBoost	<i>Accuracy</i>	97.98	99.78	98.63	99.67	100.0	93.57	95.09	98.75	99.17	99.66	
	<i>Sensitivity</i>	42.11	NULL	61.11	80.95	NULL	0.0	NULL	10.53	81.67	88.89	
	<i>Specificity</i>	100.0	99.78	100.0	100.0	100.0	100.0	95.09	100.0	100.0	99.73	
	<i>Precision</i>	100.0	NULL	100.0	100.0	NULL	0.0	NULL	100.0	100.0	66.67	
	<i>F1-score</i>	59.26	NULL	75.86	89.47	NULL	0.0	NULL	19.05	89.91	76.19	
	<i>AUC</i>	99.67	NULL	89.06	99.97	NULL	93.37	NULL	99.02	96.78	99.8	
	<i>FP/h</i>	0.0	8.02	0.0	0.0	0.0	0.0	176.8	0.0	0.0	9.76	
	<i>CPU Time</i>	0.66	0.91	1.61	0.92	0.89	0.58	0.86	0.83	0.8	0.89	
LightGBM	<i>Accuracy</i>	99.27	99.41	97.52	99.59	100.0	93.66	96.73	98.53	99.24	99.93	
	<i>Sensitivity</i>	84.21	NULL	29.63	76.19	NULL	1.45	NULL	0.0	83.33	88.89	
	<i>Specificity</i>	99.81	99.41	100.0	100.0	100.0	100.0	96.73	99.93	100.0	100.0	
	<i>Precision</i>	94.12	NULL	100.0	100.0	NULL	100.0	NULL	0.0	100.0	100.0	
	<i>F1-score</i>	88.89	NULL	45.71	86.49	NULL	2.86	NULL	0.0	90.91	94.12	
	<i>AUC</i>	97.26	NULL	89.13	98.93	NULL	93.55	NULL	98.43	94.55	99.95	
	<i>FP/h</i>	6.61	21.4	0.0	0.0	0.0	0.0	117.86	2.64	0.0	0.0	
	<i>CPU Time</i>	0.12	0.06	0.25	0.06	0.2	0.11	0.11	0.11	0.11	0.12	

Note. All metrics are in percent apart from FP/h and CPU Time.

Table S5: NHS (Leeds) performance metrics before post-processing.

Classifier	Metric	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
BRF	<i>Accuracy</i>	96.77	88.64	98.52	97.56	96.89	88.49	97.93	93.52	98.71	95.69	98.29
	<i>Sensitivity</i>	91.67	88.13	75.0	NULL	93.48	94.44	78.57	92.5	90.62	95.45	90.11
	<i>Specificity</i>	96.89	88.66	100.0	97.56	97.16	88.31	99.68	93.61	99.64	95.7	99.51
	<i>Precision</i>	42.31	17.53	100.0	NULL	71.67	20.24	95.65	54.41	96.67	44.68	96.47
	<i>F1-score</i>	57.89	29.24	85.71	NULL	81.13	33.33	86.27	68.52	93.55	60.87	93.18
	<i>AUC</i>	99.41	94.67	87.5	NULL	96.92	98.9	96.6	95.64	97.57	98.93	99.66
	<i>FP/h</i>	109.09	397.43	0.0	87.8	95.03	408.12	10.67	212.57	11.58	149.28	15.38
	<i>CPU Time</i>	0.23	1.78	0.0	0.12	0.03	0.11	0.05	0.17	0.02	0.08	0.2
RUSBoost	<i>Accuracy</i>	98.59	95.81	98.67	99.88	98.76	95.6	98.07	95.81	98.55	99.68	96.3
	<i>Sensitivity</i>	91.67	76.56	77.5	NULL	82.61	88.89	82.14	60.0	89.06	90.91	71.43
	<i>Specificity</i>	98.76	96.34	100.0	99.88	100.0	95.81	99.52	98.76	99.64	100.0	100.0
	<i>Precision</i>	64.71	36.36	100.0	NULL	100.0	40.0	93.88	80.0	96.61	100.0	100.0
	<i>F1-score</i>	75.86	49.3	87.32	NULL	90.48	55.17	87.62	68.57	92.68	95.24	83.33
	<i>AUC</i>	99.72	93.29	97.44	NULL	96.66	98.1	99.09	96.0	98.11	98.84	96.81
	<i>FP/h</i>	43.64	128.41	0.0	4.39	0.0	146.19	16.0	41.14	11.58	0.0	0.0
	<i>CPU Time</i>	0.14	8.34	0.14	0.23	0.17	0.17	0.23	0.14	0.17	0.12	0.16
LightGBM	<i>Accuracy</i>	98.38	97.89	98.67	99.51	98.76	96.95	97.78	97.14	98.87	99.36	97.15
	<i>Sensitivity</i>	75.0	83.65	77.5	NULL	82.61	88.89	78.57	77.5	89.06	90.91	78.02
	<i>Specificity</i>	98.96	98.28	100.0	99.51	100.0	97.21	99.52	98.76	100.0	99.67	100.0
	<i>Precision</i>	64.29	57.06	100.0	NULL	100.0	50.0	93.62	83.78	100.0	90.91	100.0
	<i>F1-score</i>	69.23	67.84	87.32	NULL	90.48	64.0	85.44	80.52	94.21	90.91	87.65
	<i>AUC</i>	96.95	95.28	98.54	NULL	97.25	98.83	98.83	96.44	98.22	99.23	99.28
	<i>FP/h</i>	36.36	60.32	0.0	17.56	0.0	97.46	16.0	41.14	0.0	11.48	0.0
	<i>CPU Time</i>	0.0	0.36	0.02	0.02	0.02	0.02	0.02	0.02	0.0	0.0	0.02
Classifier	Metric	P12	P13	P14	P15	P16						
BRF	<i>Accuracy</i>	96.06	89.67	92.88	96.43	95.33						
	<i>Sensitivity</i>	78.57	93.02	75.0	90.15	66.67						
	<i>Specificity</i>	96.99	89.42	93.12	96.77	96.53						
	<i>Precision</i>	57.89	40.0	12.59	60.41	44.44						
	<i>F1-score</i>	66.67	55.94	21.56	72.34	53.33						
	<i>AUC</i>	94.04	97.08	90.83	97.0	96.54						
	<i>FP/h</i>	103.04	354.1	244.57	110.2	120.0						
	<i>CPU Time</i>	0.25	0.23	0.09	0.91	0.2						
RUSBoost	<i>Accuracy</i>	98.03	94.59	98.15	98.67	98.33						
	<i>Sensitivity</i>	71.43	79.07	41.67	78.03	62.5						
	<i>Specificity</i>	99.44	95.77	98.9	99.79	99.83						
	<i>Precision</i>	86.96	58.62	33.33	95.37	93.75						
	<i>F1-score</i>	78.43	67.33	37.04	85.83	75.0						
	<i>AUC</i>	92.34	95.21	88.21	95.55	98.71						
	<i>FP/h</i>	19.32	141.64	39.13	7.06	6.0						
	<i>CPU Time</i>	0.14	0.17	0.62	0.84	0.14						
LightGBM	<i>Accuracy</i>	98.03	95.74	98.04	98.78	98.67						
	<i>Sensitivity</i>	71.43	48.84	37.5	80.3	75.0						
	<i>Specificity</i>	99.44	99.29	98.84	99.79	99.65						
	<i>Precision</i>	86.96	84.0	30.0	95.5	90.0						
	<i>F1-score</i>	78.43	61.76	33.33	87.24	81.82						
	<i>AUC</i>	93.81	96.65	90.19	96.69	98.26						
	<i>FP/h</i>	19.32	23.61	41.09	7.06	12.0						
	<i>CPU Time</i>	0.02	0.02	0.03	0.03	0.02						

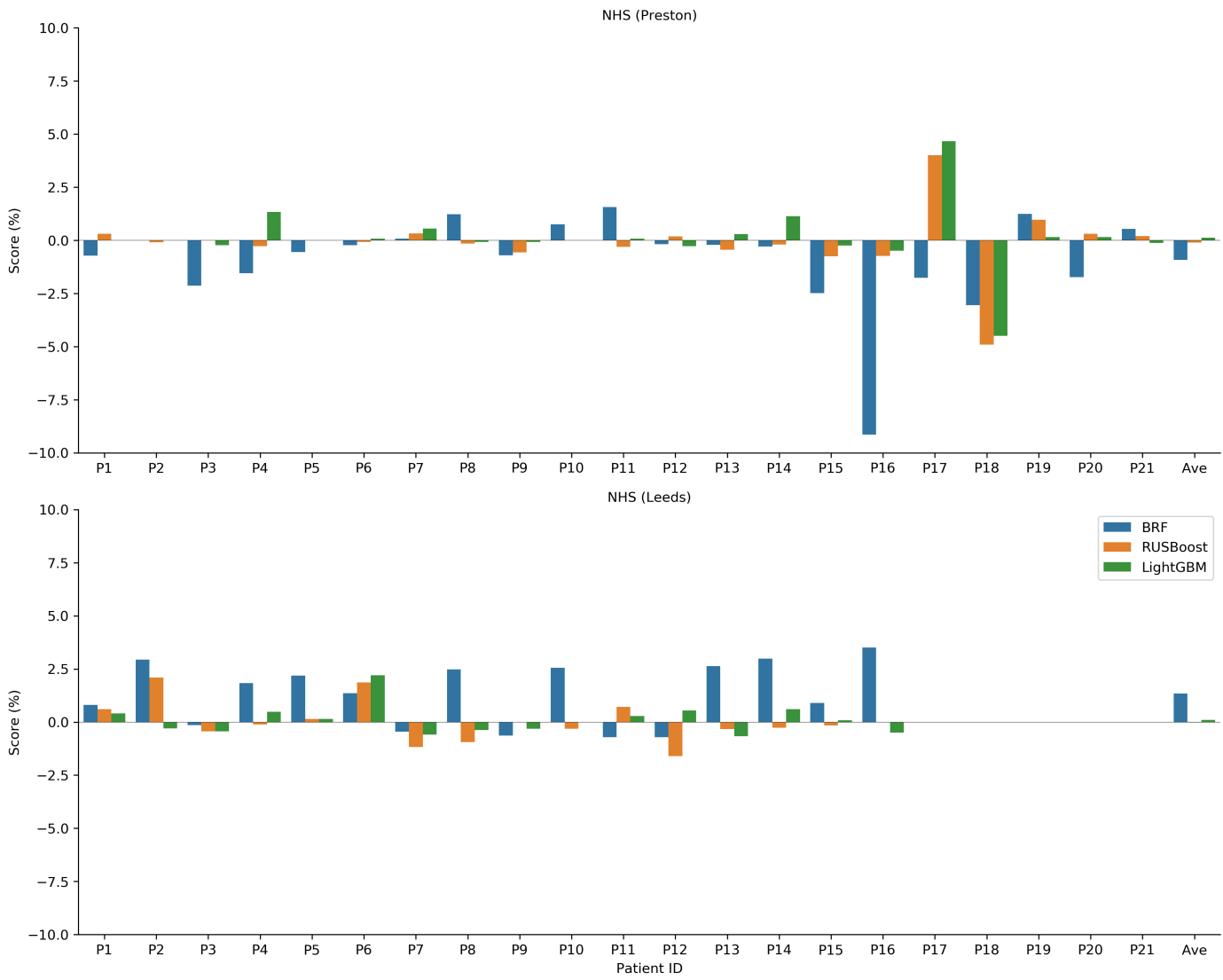
Note. All metrics are in percent apart from FP/h and CPU Time.

Table S6: NHS (Combined) performance metrics before post-processing.

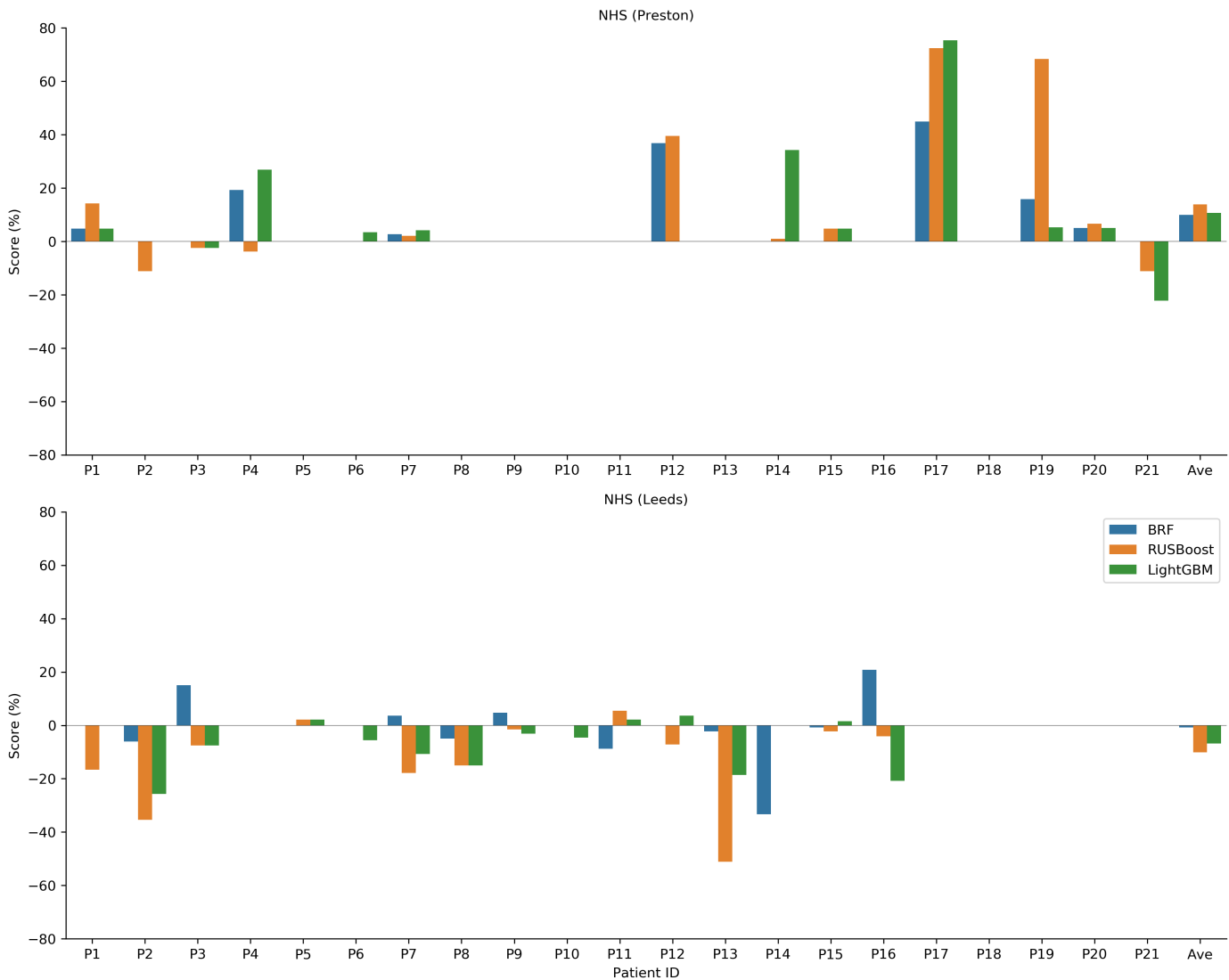
Classifier	Metric	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
BRF	<i>Accuracy</i>	97.94	98.64	96.18	96.75	98.87	98.34	98.03	96.55	94.73	99.5	94.48
	<i>Sensitivity</i>	90.48	88.89	87.8	96.15	NULL	79.31	89.42	NULL	NULL	NULL	NULL
	<i>Specificity</i>	98.07	98.72	96.45	96.78	98.87	98.77	99.54	96.55	94.73	99.5	94.48
	<i>Precision</i>	44.19	34.78	44.44	63.56	NULL	58.97	97.13	NULL	NULL	NULL	NULL
	<i>F1-score</i>	59.38	50.0	59.02	76.53	NULL	67.65	93.11	NULL	NULL	NULL	NULL
	<i>AUC</i>	97.27	97.12	95.81	98.42	NULL	95.41	97.47	NULL	NULL	NULL	NULL
	<i>FP/h</i>	68.3	45.8	123.66	109.4	40.81	43.47	14.2	124.24	189.88	17.97	198.76
	<i>CPU Time</i>	0.52	0.22	0.05	0.02	0.27	0.11	0.34	0.47	0.39	0.47	0.16
RUSBoost	<i>Accuracy</i>	99.68	99.83	99.01	98.94	100.0	99.32	96.77	99.84	99.43	99.83	99.46
	<i>Sensitivity</i>	80.95	77.78	82.93	84.62	NULL	75.86	78.31	NULL	NULL	NULL	NULL
	<i>Specificity</i>	100.0	100.0	99.53	99.78	100.0	99.85	100.0	99.84	99.43	99.83	99.46
	<i>Precision</i>	100.0	100.0	85.0	95.65	NULL	91.67	100.0	NULL	NULL	NULL	NULL
	<i>F1-score</i>	89.47	87.5	83.95	89.8	NULL	83.02	87.83	NULL	NULL	NULL	NULL
	<i>AUC</i>	97.01	97.71	93.42	98.08	NULL	90.19	96.29	NULL	NULL	NULL	NULL
	<i>FP/h</i>	0.0	0.0	16.49	7.63	0.0	5.43	0.0	5.92	20.53	5.99	19.6
	<i>CPU Time</i>	0.41	0.33	0.34	0.38	0.44	0.38	0.36	0.42	0.45	0.39	0.36
LightGBM	<i>Accuracy</i>	99.68	99.83	99.08	99.36	100.0	99.4	97.95	99.92	99.93	99.83	99.92
	<i>Sensitivity</i>	85.71	77.78	82.93	91.03	NULL	79.31	86.77	NULL	NULL	NULL	NULL
	<i>Specificity</i>	99.92	100.0	99.61	99.85	100.0	99.85	99.91	99.92	99.93	99.83	99.92
	<i>Precision</i>	94.74	100.0	87.18	97.26	NULL	92.0	99.39	NULL	NULL	NULL	NULL
	<i>F1-score</i>	90.0	87.5	85.0	94.04	NULL	85.19	92.66	NULL	NULL	NULL	NULL
	<i>AUC</i>	98.07	95.0	95.37	98.16	NULL	94.24	96.63	NULL	NULL	NULL	NULL
	<i>FP/h</i>	2.85	0.0	13.74	5.09	0.0	5.43	2.84	2.96	2.57	5.99	2.8
	<i>CPU Time</i>	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Classifier	Metric	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22
BRF	<i>Accuracy</i>	96.88	97.4	98.04	96.77	86.04	93.94	85.2	99.12	97.73	97.49	97.58
	<i>Sensitivity</i>	94.74	NULL	79.63	95.24	NULL	94.2	NULL	36.84	93.33	100.0	91.67
	<i>Specificity</i>	96.96	97.4	98.71	96.8	86.04	93.92	85.2	100.0	97.94	97.48	97.72
	<i>Precision</i>	52.94	NULL	69.35	34.48	NULL	51.59	NULL	100.0	68.29	19.57	50.0
	<i>F1-score</i>	67.92	NULL	74.14	50.63	NULL	66.67	NULL	53.85	78.87	32.73	64.71
	<i>AUC</i>	98.4	NULL	91.43	99.79	NULL	97.28	NULL	99.24	96.32	99.8	99.0
	<i>FP/h</i>	105.79	93.61	44.69	113.34	502.53	204.66	532.95	0.0	70.91	90.24	80.0
	<i>CPU Time</i>	0.41	0.58	0.09	0.11	0.45	0.16	0.17	0.55	0.03	0.33	0.03
RUSBoost	<i>Accuracy</i>	98.16	99.33	98.43	98.92	99.27	97.58	90.18	99.71	99.47	99.86	99.19
	<i>Sensitivity</i>	81.58	NULL	62.04	85.71	NULL	72.46	NULL	78.95	88.33	77.78	75.0
	<i>Specificity</i>	98.76	99.33	99.76	99.16	99.27	99.3	90.18	100.0	100.0	100.0	99.79
	<i>Precision</i>	70.45	NULL	90.54	64.29	NULL	87.72	NULL	100.0	100.0	100.0	90.0
	<i>F1-score</i>	75.61	NULL	73.63	73.47	NULL	79.37	NULL	88.24	93.81	87.5	81.82
	<i>AUC</i>	98.83	NULL	88.27	99.1	NULL	96.37	NULL	99.47	95.49	99.61	99.38
	<i>FP/h</i>	42.98	24.07	8.23	29.83	26.45	23.49	353.59	0.0	0.0	0.0	7.27
	<i>CPU Time</i>	0.39	0.48	0.98	0.42	0.42	0.34	0.48	0.45	0.45	0.45	0.12
LightGBM	<i>Accuracy</i>	98.99	99.7	98.66	99.34	99.51	98.32	92.24	98.68	99.39	99.8	98.79
	<i>Sensitivity</i>	84.21	NULL	63.89	80.95	NULL	76.81	NULL	5.26	88.33	66.67	75.0
	<i>Specificity</i>	99.52	99.7	99.93	99.66	99.51	99.8	92.24	100.0	99.92	100.0	99.38
	<i>Precision</i>	86.49	NULL	97.18	80.95	NULL	96.36	NULL	100.0	98.15	100.0	75.0
	<i>F1-score</i>	85.33	NULL	77.09	80.95	NULL	85.48	NULL	10.0	92.98	80.0	75.0
	<i>AUC</i>	98.96	NULL	88.06	99.45	NULL	98.04	NULL	99.88	96.63	99.38	99.29
	<i>FP/h</i>	16.53	10.7	2.35	11.93	17.63	6.71	279.29	0.0	2.73	0.0	21.82
	<i>CPU Time</i>	0.03	0.05	0.08	0.03	0.02	0.03	0.05	0.02	0.02	0.03	0.02

Classifier	Metric	P23	P24	P25	P26	P27	P28	P29	P30	P31	P32	P33
BRF	<i>Accuracy</i>	91.58	98.38	99.39	99.07	89.85	97.48	96.0	98.07	98.25	97.58	95.35
	<i>Sensitivity</i>	82.05	90.0	NULL	93.48	94.44	82.14	87.5	95.31	95.45	81.32	78.57
	<i>Specificity</i>	91.84	98.9	99.39	99.5	89.7	98.87	96.7	98.39	98.35	100.0	96.23
	<i>Precision</i>	21.56	83.72	NULL	93.48	22.37	86.79	68.63	87.14	67.74	100.0	52.38
	<i>F1-score</i>	34.15	86.75	NULL	93.48	36.17	84.4	76.92	91.04	79.25	89.7	62.86
	<i>AUC</i>	92.96	97.32	NULL	97.08	99.19	99.28	94.97	98.56	98.81	99.27	93.09
	<i>FP/h</i>	286.08	37.22	21.95	16.77	359.39	37.33	109.71	52.09	57.42	0.0	128.8
	<i>CPU Time</i>	4.78	0.09	0.17	0.12	0.14	0.31	0.09	0.17	0.17	0.11	0.02
RUSBoost	<i>Accuracy</i>	97.91	98.23	99.76	98.91	97.46	96.89	94.86	98.55	99.36	97.01	96.42
	<i>Sensitivity</i>	41.1	70.0	NULL	84.78	88.89	64.29	45.0	87.5	90.91	76.92	64.29
	<i>Specificity</i>	99.46	100.0	99.76	100.0	97.73	99.84	98.97	99.82	99.67	100.0	98.12
	<i>Precision</i>	67.62	100.0	NULL	100.0	55.17	97.3	78.26	98.25	90.91	100.0	64.29
	<i>F1-score</i>	51.13	82.35	NULL	91.76	68.09	77.42	57.14	92.56	90.91	86.96	64.29
	<i>AUC</i>	92.31	93.73	NULL	96.36	98.47	98.45	95.16	99.17	97.38	99.09	89.25
	<i>FP/h</i>	18.86	0.0	8.78	0.0	79.19	5.33	34.29	5.79	11.48	0.0	64.4
	<i>CPU Time</i>	6.89	0.17	0.28	0.14	0.19	0.17	0.12	0.14	0.17	0.16	0.12
LightGBM	<i>Accuracy</i>	97.59	98.23	100.0	98.91	99.15	97.19	96.76	98.55	99.36	97.44	98.57
	<i>Sensitivity</i>	57.89	70.0	NULL	84.78	83.33	67.86	62.5	85.94	86.36	80.22	75.0
	<i>Specificity</i>	98.67	100.0	100.0	100.0	99.65	99.84	99.59	100.0	99.83	100.0	99.81
	<i>Precision</i>	54.42	100.0	NULL	100.0	88.24	97.44	92.59	100.0	95.0	100.0	95.45
	<i>F1-score</i>	56.1	82.35	NULL	91.76	85.71	80.0	74.63	92.44	90.48	89.02	84.0
	<i>AUC</i>	94.24	97.35	NULL	95.74	98.82	99.31	96.07	98.79	97.24	99.49	91.91
	<i>FP/h</i>	46.45	0.0	0.0	0.0	12.18	5.33	13.71	0.0	5.74	0.0	6.44
	<i>CPU Time</i>	0.59	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.0	0.0
Classifier	Metric	P34	P35	P36	P37							
BRF	<i>Accuracy</i>	92.3	95.87	97.33	98.83							
	<i>Sensitivity</i>	90.7	41.67	89.39	87.5							
	<i>Specificity</i>	92.42	96.59	97.76	99.31							
	<i>Precision</i>	47.56	13.89	68.6	84.0							
	<i>F1-score</i>	62.4	20.83	77.63	85.71							
	<i>AUC</i>	97.82	87.98	96.34	98.38							
	<i>FP/h</i>	253.77	121.3	76.3	24.0							
	<i>CPU Time</i>	0.06	0.53	0.03	0.14							
RUSBoost	<i>Accuracy</i>	94.26	97.88	98.51	98.33							
	<i>Sensitivity</i>	27.91	41.67	75.76	58.33							
	<i>Specificity</i>	99.29	98.62	99.75	100.0							
	<i>Precision</i>	75.0	28.57	94.34	100.0							
	<i>F1-score</i>	40.68	33.9	84.03	73.68							
	<i>AUC</i>	93.71	84.49	96.66	97.09							
	<i>FP/h</i>	23.61	48.91	8.48	0.0							
	<i>CPU Time</i>	0.14	0.64	0.72	0.14							
LightGBM	<i>Accuracy</i>	95.08	98.64	98.86	98.17							
	<i>Sensitivity</i>	30.23	37.5	81.82	54.17							
	<i>Specificity</i>	100.0	99.45	99.79	100.0							
	<i>Precision</i>	100.0	47.37	95.58	100.0							
	<i>F1-score</i>	46.43	41.86	88.16	70.27							
	<i>AUC</i>	95.31	88.09	96.29	98.2							
	<i>FP/h</i>	0.0	19.57	7.06	0.0							
	<i>CPU Time</i>	0.02	0.03	0.05	0.02							

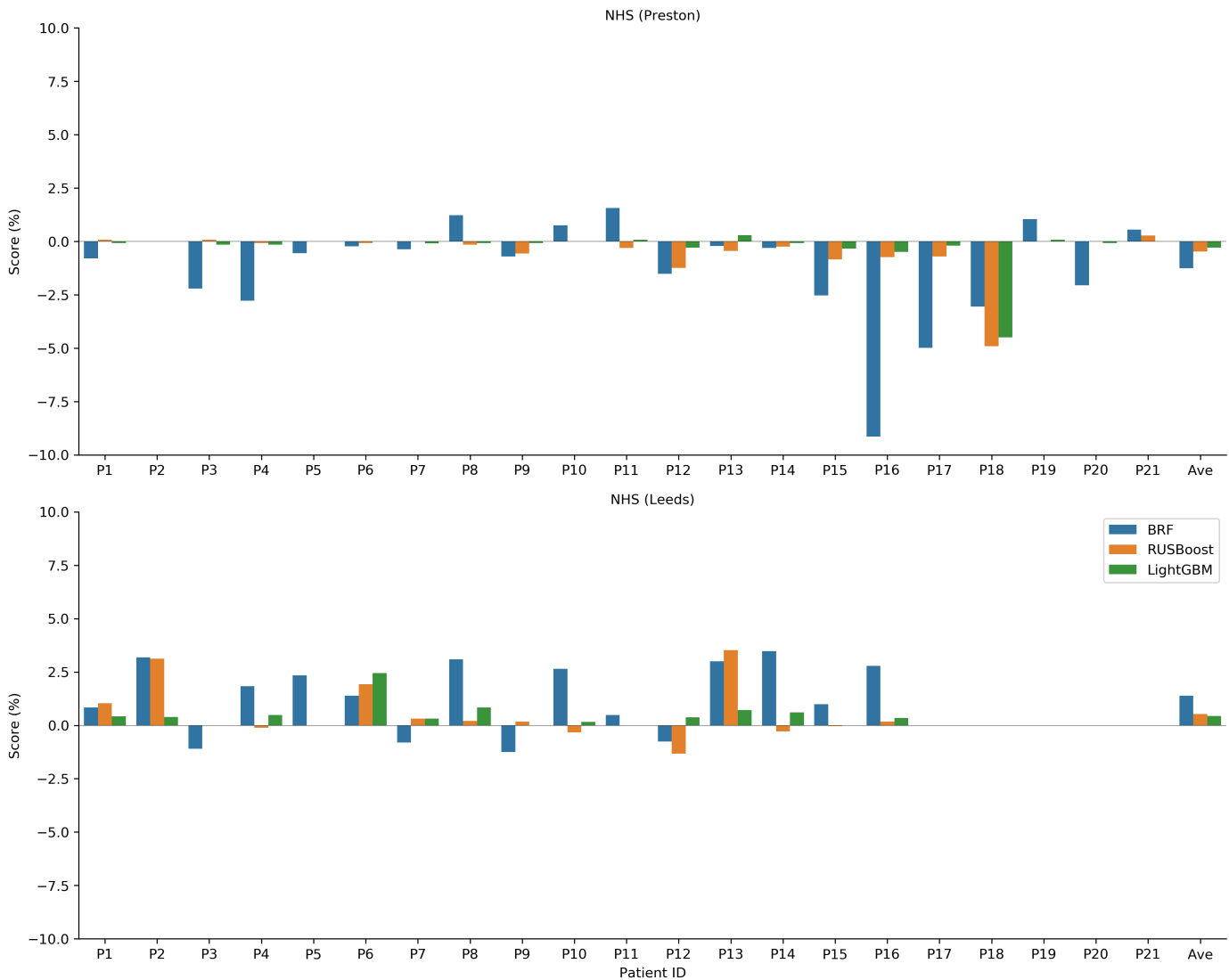
Note. All metrics are in percent apart from FP/h and CPU Time. P1 - P21 (Preston), P22 - P37 (Leeds).



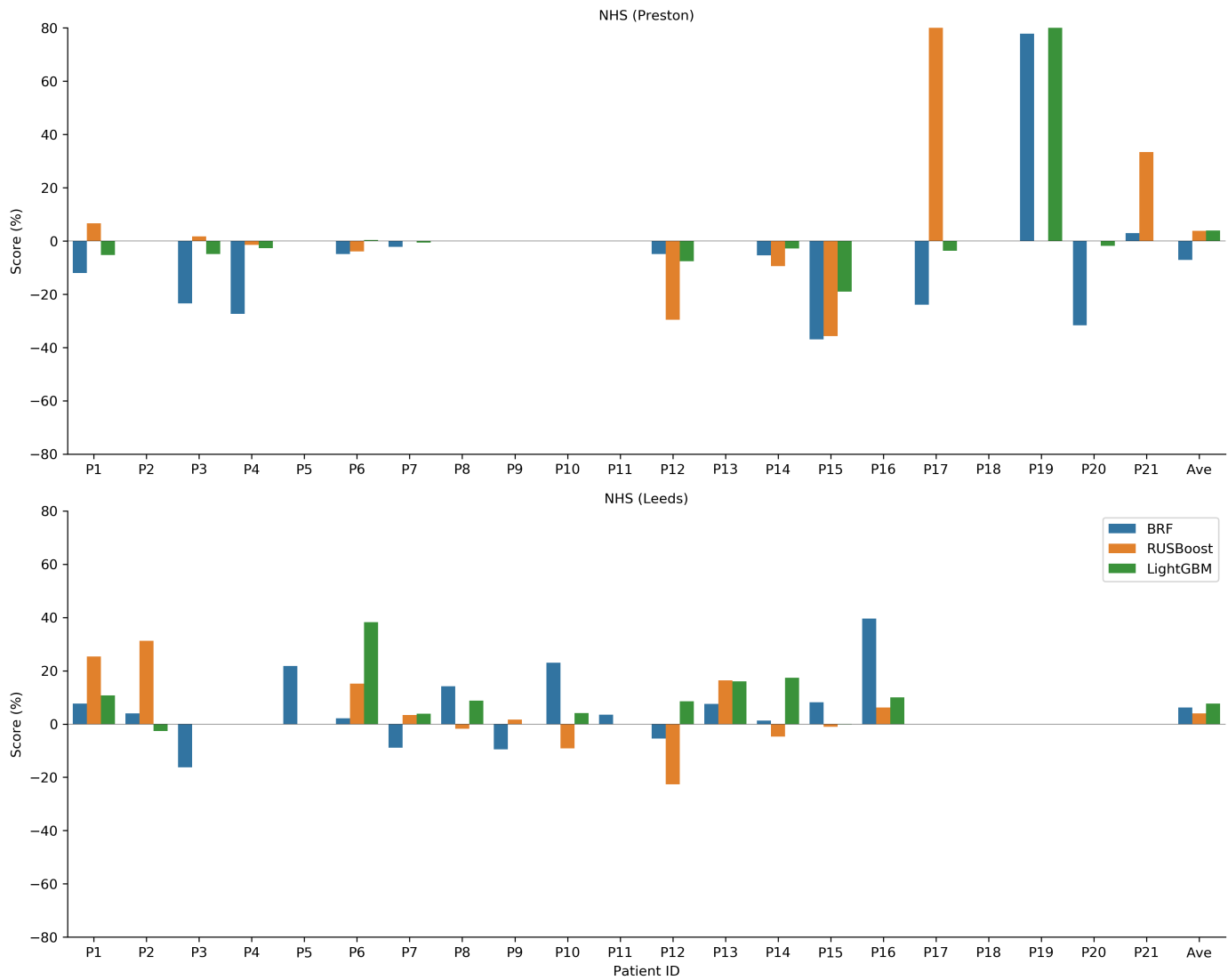
(a) Accuracy



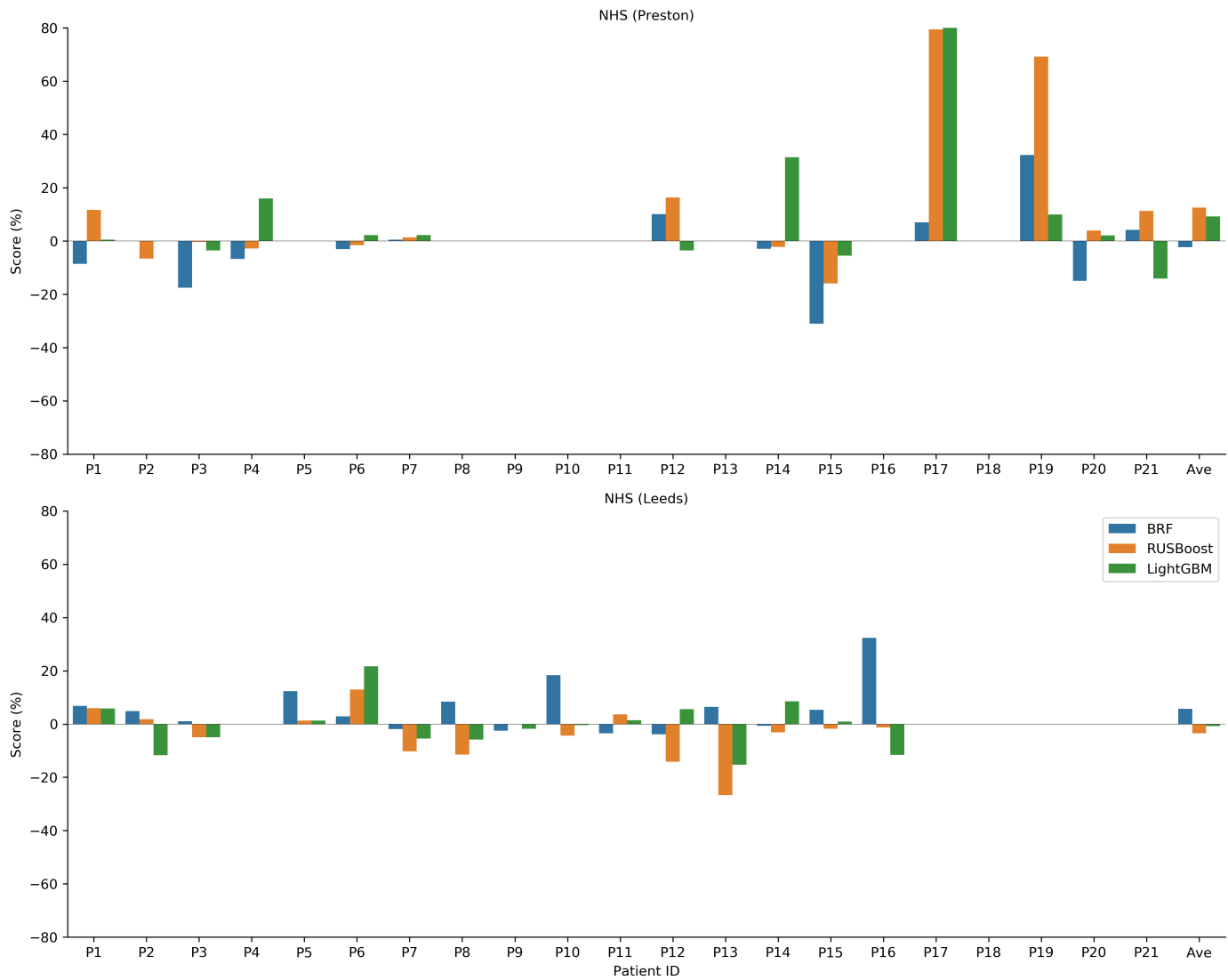
(b) Sensitivity



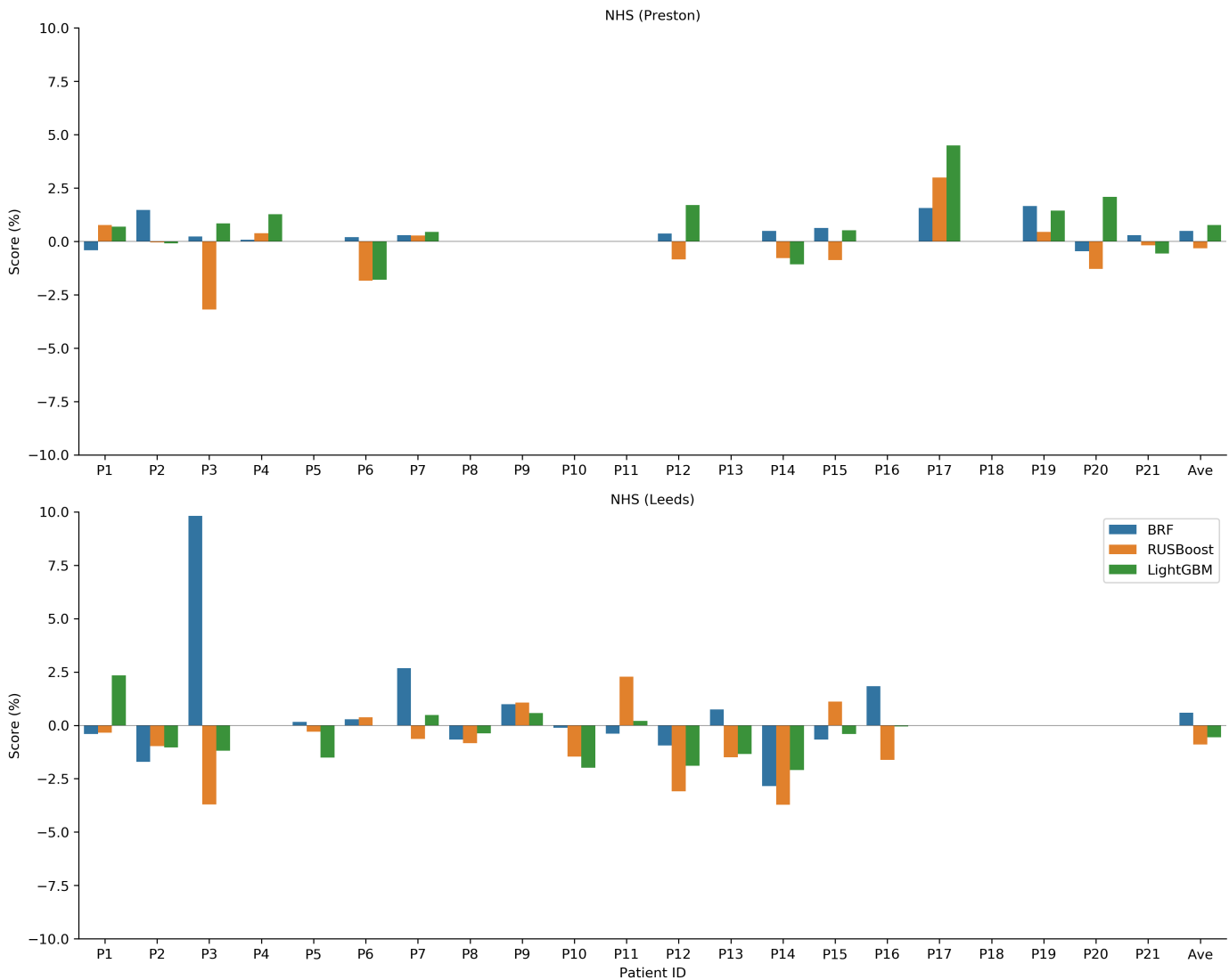
(c) Specificity



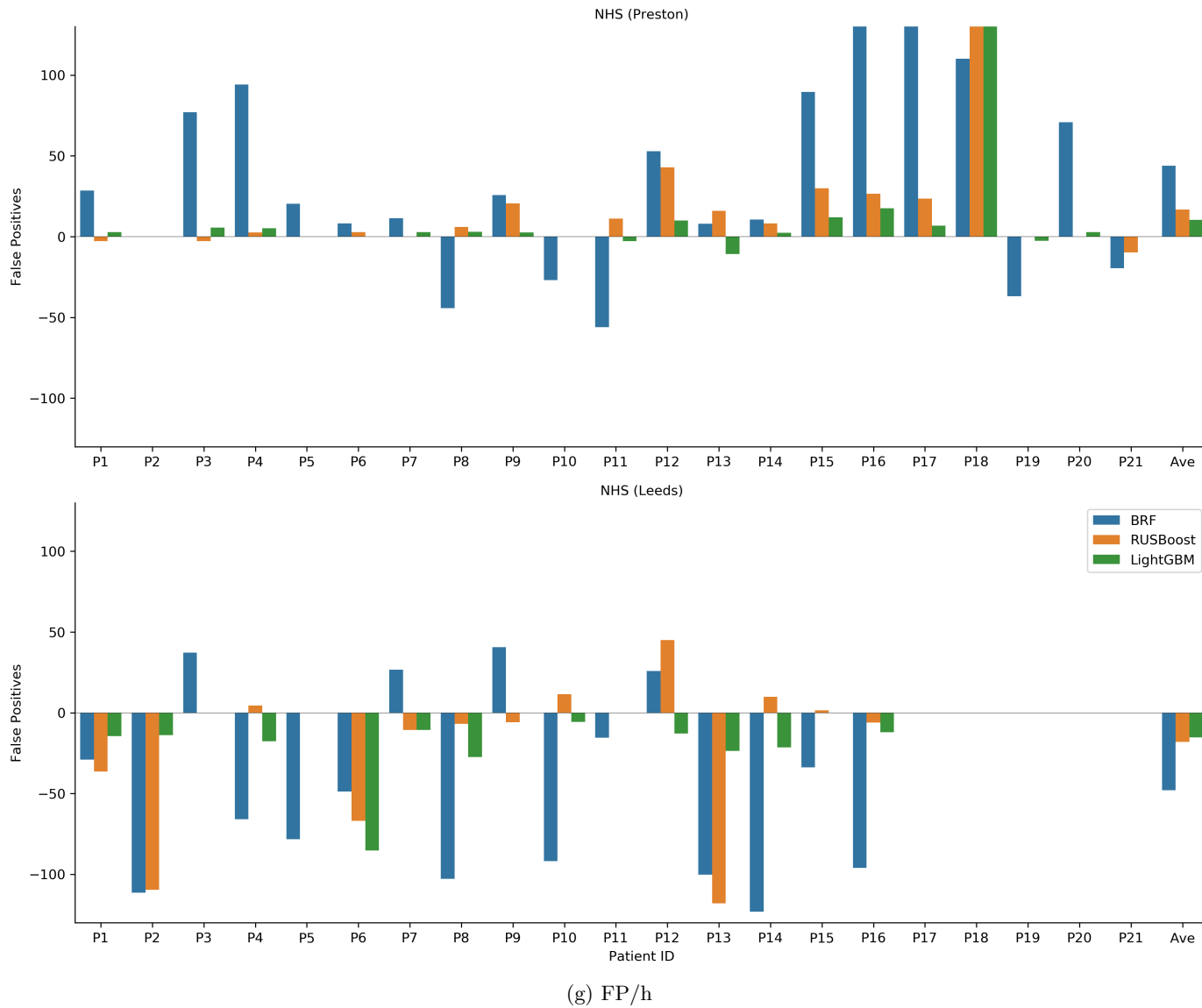
(d) Precision



(e) F1-score



(f) AUC



(g) FP/h

Figure S4: Change in performance from when models were trained on NHS datasets separately to when they were combined.

Chapter 5

As seen in table S6, performance was particularly poor for all models on record P7. This record also had the poorest performance in chapter 4, as models mark seizures at the beginning of the record where there is a prolonged period of relative inactivity, and do not fully mark seizures where present. However, only one other record (P2) had any false positives for optimal LightGBM models, demonstrating the good specificity of this model.

Table S6: TUH (Absence) performance before post-processing.

Classifier	Metric	00000675	00001113	00001413	00001795	00001984	00002448	00002657	00003053	00003281	00003306	00003635
LightGBM	Accuracy	98.3	99.13	99.49	98.64	96.95	98.6	95.17	99.59	99.46	99.64	98.79
	Sensitivity	85.77	83.49	82.29	76.6	59.8	79.07	89.47	64.71	86.96	87.8	81.82
	Specificity	99.77	99.77	99.97	99.62	100.0	99.85	95.52	100.0	99.69	100.0	99.15
	Precision	97.75	93.81	98.75	90.0	100.0	97.14	55.28	100.0	83.33	100.0	67.5
	F1-score	91.37	88.35	89.77	82.76	74.85	87.18	68.34	78.57	85.11	93.51	73.97
	AUC	92.77	91.63	91.13	88.11	79.9	89.46	92.5	82.35	93.32	93.9	90.48
	FP/h	7.47	7.87	1.03	13.02	0.0	5.04	151.78	0.0	11.13	0.0	29.92
	Prediction Time	0.03	0.09	0.11	0.02	0.02	0.03	0.03	0.02	0.01	0.01	0.02
	Accuracy	97.97	98.83	99.43	98.73	96.88	98.13	94.86	99.38	99.54	99.57	98.66
	Sensitivity	83.0	86.24	81.25	74.47	63.73	74.42	88.82	76.47	86.96	92.68	72.73
MLP	Specificity	99.72	99.36	99.94	99.81	99.6	99.65	95.24	99.65	99.76	99.78	99.22
	Precision	97.22	84.68	97.5	94.59	92.86	93.2	53.57	72.22	86.96	92.68	66.67
	F1-score	89.55	85.45	88.64	83.33	75.58	82.76	66.83	74.29	86.96	92.68	69.57
	AUC	91.36	92.8	90.6	87.14	81.66	87.04	92.03	88.06	93.36	96.23	85.97
	FP/h	8.96	22.3	2.05	6.51	13.38	11.75	161.44	12.41	8.35	7.76	27.62
	Prediction Time	0.14	0.22	0.3	0.11	0.11	0.45	0.16	0.12	0.08	0.08	0.1
	Accuracy	98.47	98.94	99.39	98.43	97.61	97.23	95.29	99.72	99.14	99.86	98.42
	Sensitivity	86.64	82.88	79.41	67.35	66.34	76.09	88.89	83.33	87.5	97.37	81.25
	Specificity	99.82	99.62	99.95	99.74	100.0	98.67	95.68	99.93	99.36	99.93	98.77
	Precision	98.17	90.2	97.59	91.67	100.0	79.55	56.2	93.75	72.41	97.37	57.78
CNN RNN	F1-score	92.04	86.38	87.57	77.65	79.76	77.78	68.86	88.24	79.25	97.37	67.53
	AUC	93.23	91.25	89.68	83.54	83.17	87.38	92.29	91.63	93.43	98.65	90.01
	FP/h	5.95	13.13	1.92	8.93	0.0	44.94	146.26	2.51	22.48	2.59	43.32
	Prediction Time	1.74	1.9	2.59	1.2	1.35	2.12	1.86	1.29	0.71	0.76	1.13
	Accuracy	98.14	98.76	98.93	98.43	97.12	97.36	94.56	99.51	99.3	99.86	98.29
	Sensitivity	82.19	74.77	66.67	65.31	63.37	74.64	81.05	83.33	75.0	94.74	78.12
	Specificity	99.95	99.77	99.84	99.83	99.7	98.91	95.4	99.72	99.76	100.0	98.71
	Precision	99.51	93.26	91.89	94.12	94.12	82.4	52.32	78.95	85.71	100.0	55.56
	F1-score	90.02	83.0	77.27	77.11	75.74	78.33	63.59	81.08	80.0	97.3	64.94
	AUC	91.07	87.27	83.25	82.57	81.53	86.78	88.22	91.53	87.38	97.37	88.42
CNN1D	FP/h	1.49	7.88	5.77	5.95	10.12	36.62	155.92	10.03	8.43	0.0	45.6
	Prediction Time	0.35	0.51	0.7	1.81	0.29	3.23	0.35	0.31	0.18	0.16	0.21
	Accuracy	97.81	96.02	99.65	98.43	97.54	97.18	94.94	99.02	98.59	99.86	99.05
	Sensitivity	79.35	95.5	88.24	67.35	80.2	60.14	86.27	88.89	83.33	94.74	62.5
	Specificity	99.91	96.05	99.97	99.74	98.87	99.7	95.48	99.15	98.89	100.0	99.81
	Precision	98.99	50.48	98.9	91.67	84.38	93.26	54.32	57.14	58.82	100.0	86.96
	F1-score	88.09	66.04	93.26	77.65	82.23	73.13	66.67	69.57	68.97	97.3	72.73
	AUC	89.63	95.77	94.1	83.54	89.53	79.92	90.88	94.02	91.11	97.37	81.15
	FP/h	2.97	136.59	0.96	8.93	37.95	9.99	153.16	30.1	39.34	0.0	6.84
	Prediction Time	1.05	2.29	3.1	0.65	0.61	1.06	1.25	0.69	1.68	1.68	0.85

Note. All metrics are in percent apart from FP/h and prediction time.

Table S7: TUH (Absence) performance after post-processing.

Classifier	Metric	00000675	00001113	00001413	00001795	00001984	00002448	00002657	00003053	00003281	00003306	00003635
LightGBM	Accuracy	97.64	99.16	99.43	99.01	96.95	98.65	96.13	99.59	99.77	99.64	99.3
	Sensitivity	77.47	81.65	79.17	76.6	59.8	77.52	89.47	64.71	86.96	87.8	66.67
	Specificity	100.0	99.89	100.0	100.0	100.0	100.0	96.54	100.0	100.0	100.0	100.0
	Precision	100.0	96.74	100.0	100.0	100.0	100.0	61.54	100.0	100.0	100.0	100.0
	F1-score	87.31	88.56	88.37	86.75	74.85	87.34	72.92	78.57	93.02	93.51	80.0
	AUC	88.74	90.77	89.58	88.3	79.9	88.76	93.01	82.35	93.48	93.9	83.33
	FP/h	0.0	3.93	0.0	0.0	0.0	0.0	117.29	0.0	0.0	0.0	0.0
	Prediction Time	0.03	0.09	0.11	0.02	0.02	0.03	0.03	0.02	0.01	0.01	0.02
	Accuracy	97.84	99.2	99.26	98.82	97.1	98.09	96.44	99.72	99.77	99.71	99.17
	Sensitivity	79.84	84.4	73.96	72.34	61.76	70.54	88.16	76.47	86.96	92.68	60.61
MLP	Specificity	99.95	99.81	99.97	100.0	100.0	99.85	96.95	100.0	100.0	99.93	100.0
	Precision	99.51	94.85	98.61	100.0	100.0	96.81	64.11	100.0	100.0	97.44	100.0
	F1-score	88.6	89.32	84.52	83.95	76.36	81.61	74.24	86.67	93.02	95.0	75.47
	AUC	89.9	92.11	86.96	86.17	80.88	85.2	92.55	88.24	93.48	96.3	80.3
	FP/h	1.49	6.56	1.03	0.0	0.0	5.04	103.49	0.0	0.0	2.59	0.0
	Prediction Time	0.14	0.22	0.3	0.11	0.11	0.45	0.16	0.12	0.08	0.08	0.1
	Accuracy	98.51	99.2	99.36	98.51	97.54	97.78	96.55	99.79	99.45	99.93	99.24
	Sensitivity	86.64	81.98	76.47	63.27	65.35	73.91	88.89	83.33	83.33	97.37	81.25
	Specificity	99.86	99.92	100.0	100.0	100.0	99.41	97.03	100.0	99.76	100.0	99.61
	Precision	98.62	97.85	100.0	100.0	100.0	89.47	65.07	100.0	86.96	100.0	81.25
CNN RNN	F1-score	92.24	89.22	86.67	77.5	79.04	80.95	75.14	90.91	85.11	98.67	81.25
	AUC	93.25	90.95	88.24	81.63	82.67	86.66	92.96	91.67	91.55	98.68	90.43
	FP/h	4.46	2.63	0.0	0.0	0.0	19.97	100.73	0.0	8.43	0.0	13.68
	Prediction Time	1.74	1.9	2.59	1.2	1.35	2.12	1.86	1.29	0.71	0.76	1.13
	Accuracy	97.23	98.87	98.77	98.51	97.33	97.97	95.75	99.79	99.53	99.86	99.05
	Sensitivity	72.87	72.07	54.9	63.27	62.38	71.01	80.39	83.33	75.0	94.74	53.12
	Specificity	100.0	100.0	100.0	100.0	100.0	99.8	96.7	100.0	100.0	100.0	100.0
	Precision	100.0	100.0	100.0	100.0	100.0	96.08	60.29	100.0	100.0	100.0	100.0
	F1-score	84.31	83.77	70.89	77.5	76.83	81.67	68.91	90.91	85.71	97.3	69.39
	AUC	86.44	86.04	77.45	81.63	81.19	85.41	88.55	91.67	87.5	97.37	76.56
CNN1D	FP/h	0.0	0.0	0.0	0.0	0.0	6.66	111.77	0.0	0.0	0.0	0.0
	Prediction Time	0.35	0.51	0.7	1.81	0.29	3.23	0.35	0.31	0.18	0.16	0.21
	Accuracy	97.48	97.77	99.68	98.68	98.03	96.9	95.9	99.65	99.38	99.86	99.05
	Sensitivity	75.3	95.5	88.24	67.35	80.2	51.45	86.27	88.89	83.33	94.74	53.12
	Specificity	100.0	97.87	100.0	100.0	100.0	99.39	100.0	96.5	99.79	99.68	100.0
	Precision	100.0	65.43	100.0	100.0	91.01	100.0	60.55	84.21	83.33	100.0	100.0
	F1-score	85.91	77.66	93.75	80.49	85.26	67.94	71.16	86.49	83.33	97.3	69.39
	AUC	87.65	96.68	94.12	83.67	89.8	75.72	91.39	94.34	91.51	97.37	76.56
	FP/h	0.0	73.55	0.0	0.0	20.24	0.0	118.67	7.53	11.24	0.0	0.0
	Prediction Time	1.05	2.29	3.1	0.65	0.61	1.06	1.25	0.69	1.68	1.68	0.85

Note. All metrics are in percent apart from FP/h and prediction time.

Table S8: TUH (Generalised) training performance before post-processing.

Classifier	Metric	P1	P2	P4	P5	P6	P8	P11	P12	P13	P15	P16	P18	P21	P22	P24	P25	P27	P28	P29	P30
LightGBM	<i>Accuracy</i>	85.97	91.84	91.85	83.17	88.58	71.89	70.37	95.36	92.83	80.89	73.8	87.58	92	93.54	79.5	96.42	79.99	98.9	96.33	87.09
	<i>Sensitivity</i>	90.96	96.43	95.44	68.3	30.23	7.52	70.37	88.95	67.3	86.06	64.14	91.3	93.1	93.55	56.41	96.61	72.05	69.52	78.72	30.11
	<i>Specificity</i>	84.91	91.81	88.77	88.73	97.52	83.49	-	96.99	95.79	77.77	74.68	83.43	91.98	93.54	94.68	96.41	81.46	99.5	97.3	90.74
	<i>Precision</i>	56.1	7.87	87.95	69.37	65.15	7.59	100	88.33	65	70.01	18.83	86.01	15.79	87.44	87.46	54.29	41.84	73.74	61.65	17.24
	<i>F1-score</i>	69.4	14.56	91.54	68.83	41.3	7.56	82.61	88.64	66.13	77.21	29.11	88.58	27	90.39	68.59	69.51	52.94	71.57	69.15	21.92
	<i>AUC</i>	87.94	94.12	92.1	78.51	63.88	45.51	-	92.97	81.55	81.91	69.41	87.37	92.54	93.54	75.55	96.51	76.75	84.51	88.01	60.42
	<i>FP/h</i>	448.22	292.74	217.45	295.37	77.33	503.51	0	86.17	135.68	499.23	835.03	281.95	284.05	157.07	115.49	123.69	563.16	17.72	92.16	313.45
	<i>Prediction Time</i>	0.02	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.03	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.02	0.02	0.04	0.03
	<i>Prediction Time</i>	73.96	91.82	79.19	65.99	88.12	67.36	90.74	29.4	85.55	76.85	75.67	85.39	92.49	63.18	89.28	92.27	65.51	94.07	91.57	83.7
MLP	<i>Accuracy</i>	95.48	96.43	93.78	95.62	46.57	45.13	90.74	82.11	79.84	91.06	78.88	92.48	75.86	98.92	86.79	96.61	94.17	89.52	91.9	69.32
	<i>Sensitivity</i>	69.4	91.78	66.67	54.91	94.48	71.37	-	15.93	86.21	68.28	75.37	77.47	92.76	45.99	90.92	92.08	60.21	94.17	91.55	84.62
	<i>Specificity</i>	39.81	7.85	70.74	44.22	56.38	22.13	100	19.98	40.19	63.38	22.68	82.07	14.47	46.82	86.27	34.97	30.46	23.74	37.51	22.41
	<i>Precision</i>	56.19	14.52	80.64	60.47	51.01	29.69	95.15	32.14	53.47	74.74	35.23	86.97	24.31	63.56	86.53	51.35	46.04	37.52	53.28	33.87
	<i>F1-score</i>	82.44	94.11	80.22	75.27	70.53	58.25	-	49.02	83.02	79.67	77.13	84.98	84.31	72.46	88.85	94.34	77.19	91.84	91.73	76.97
	<i>AUC</i>	908.89	293.67	645.45	1181.49	172.31	873.24	0	2410.29	444.77	712.47	812.17	383.38	256.44	1313.09	197.18	273.16	1208.69	205.87	288.18	520.36
	<i>FP/h</i>	0.11	0.18	0.07	0.08	0.2	0.09	0.03	0.18	0.17	0.25	0.14	0.11	0.1	0.05	0.08	0.08	0.19	0.23	0.44	0.25
	<i>Prediction Time</i>	23.58	9.08	46.61	40.2	17.09	28.87	100	21.34	17.4	46.36	13.37	56.21	16.44	36.65	46.84	12.73	23.16	10.59	9.96	10.96
	<i>Prediction Time</i>	97.54	100	96.8	81.79	97.36	85.08	100	97.98	97.55	95.19	95.43	97.69	96.55	97.85	96.57	100	98.18	99.05	98.49	97.16
RNN	<i>Accuracy</i>	7.24	8.42	3.44	24.65	4.79	18.44	-	1.67	8.1	17.06	5.81	9.56	15.19	7.24	14.19	8.86	9.22	8.79	5.08	5.44
	<i>Sensitivity</i>	18.86	0.79	46.3	28.86	13.55	16.21	100	20.36	10.97	40.78	8.55	54.85	1.74	33.64	42.49	4.64	16.74	2.16	5.41	6.18
	<i>Specificity</i>	31.6	1.57	62.64	42.67	23.78	27.23	100	33.72	19.72	57.1	15.69	70.26	3.42	50.07	59.01	8.86	28.6	4.22	10.26	11.61
	<i>Precision</i>	52.39	54.21	50.12	53.22	51.08	51.76	-	49.83	52.82	56.12	50.62	53.63	55.87	52.54	55.38	54.43	53.7	53.92	51.78	51.3
	<i>F1-score</i>	2734.99	3273.05	1868.8	1974.44	2972.12	2476.75	0	2817.03	2964.47	1866.21	3104.67	1532.29	3006.34	2255.5	1864.85	3141.77	2755.9	3218.25	3238.76	3199.11
	<i>AUC</i>	0.34	0.64	0.18	0.31	0.88	0.27	0.04	0.47	0.76	1.46	0.55	0.44	0.41	0.13	0.28	0.3	0.86	1.07	2.12	1.19
	<i>FP/h</i>	80.37	80.09	68.68	45.01	71.73	57.56	73.17	25.37	75.89	68.62	81.23	74.85	89.96	75.22	55.8	90.49	77.31	87.01	85.51	53.71
	<i>Prediction Time</i>	86.97	86.36	74.4	77.57	83.3	22.65	73.17	93.05	80.93	81.52	55.33	82.9	89.66	82.26	56.16	96.55	54.31	94.29	94.9	54.83
	<i>Prediction Time</i>	78.91	80.04	63.76	32.84	69.95	64.04	-	8	75.3	60.89	83.61	65.8	89.96	71.83	55.57	90.22	81.59	86.86	84.99	53.64
CNN1D	<i>Accuracy</i>	47.68	3.07	63.84	30.15	29.81	10.46	100	20.61	27.55	55.57	23.75	73.16	12.21	58.4	45.35	30.43	35.4	12.71	25.85	7.04
	<i>Sensitivity</i>	61.6	5.93	68.72	43.43	43.91	14.31	84.51	33.74	41.11	66.09	33.23	77.73	21.49	68.3	50.18	46.28	42.87	22.4	40.63	12.48
	<i>Specificity</i>	82.94	83.2	69.08	55.21	76.63	43.34	-	50.52	78.11	71.2	69.47	74.35	89.81	77.05	55.87	93.39	67.95	90.57	89.94	54.23
	<i>Precision</i>	621.8	713.34	701.36	1759.94	938.03	1092.13	0	2635.71	796.71	880.07	540.08	579.45	355.81	684.82	965.57	337.09	559.03	463.55	512	1568.46
	<i>F1-score</i>	5.3	10.27	3.37	5.18	13.67	3.72	0.81	9.16	11.31	19.33	8.28	7.34	6.07	1.84	4.06	5.03	14.22	16.88	33.26	19.33
	<i>AUC</i>	93.67	95.4	94	97.78	91.06	90.98	85.82	85.3	87.75	87.03	94.41	90.19	75.93	82.49	98.04	95.48	85.73	94.05	95.87	
	<i>FP/h</i>	0	99.07	41.32	83.38	57.63	77.23	18.56	71.79	88.9	74.14	82.94	94.33	76.08	32.29	26.42	74.96	91.55	85.19	80.57	
	<i>Prediction Time</i>	100	92.27	95.47	98.54	94.11	91.3	89.68	85.7	86.13	88.24	87.98	75.9	97.05	98.98	97.53	82.71	95.53	96.23		
	<i>Prediction Time</i>	0	91.63	20.24	75.35	47.22	16.96	9.37	12.94	90.1	37.07	55.59	80.67	40.8	76.01	25.45	75.25	73.3	76.08	33.57	
LightGBM	<i>Accuracy</i>	0	95.21	27.17	79.16	51.91	27.81	12.45	21.93	89.5	49.43	66.56	86.97	53.12	45.32	25.93	75.1	81.42	80.38	47.39	
	<i>Sensitivity</i>	50	95.67	68.39	90.96	75.87	84.26	54.12	78.75	87.51	81.19	89.09	91.16	75.99	64.67	62.7	86.25	87.13	90.36	88.4	
	<i>Specificity</i>	0	150.06	158.8	49.74	194.15	306.21	351.22	500.12	206.28	387.32	159.89	282.45	712.19	82.45	36.08	80.7	409.84	137.93	132.49	
	<i>Precision</i>	0.01	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.09	0.03	0.03	0.03	
	<i>F1-score</i>	12.91	51.83	85.51	97.17	90.89	83.26	67.58	57.41	90.2	63.6	88.19	78.97	57.75	81.01	94.3	92.38	72.76	87.93	88.26	
	<i>AUC</i>	89.47	96.95	71.9	88.05	94.92	59.41	29.94	76.07	93.94	93.97	82.94	95.97	93.37	65.2	81.13	91.67	94.61	91.59	85.14	
	<i>FP/h</i>	7.73	13.31	85.89	97.66	90.52	83.8	69.74	56.86	84.9	60.76	88.57	69.93	49.97	85.59	94.48	92.45	61.43	87.32	88.33	
	<i>Prediction Time</i>	6.15	48.84	12.43	66.73	47.78	7.78	5.38	4.96	89.83	18.29	34.26	62.93	28.95	56.75	16.17	54.85	55.98	54.66	14.71	
	<i>Prediction Time</i>	11.51	64.96	21.19	75.92	63.56	13.76	9.12	9.32	91.84	30.62	48.49	76.02	44.2	60.69	26.96	68.64	70.34	68.46	25.08	
MLP	<i>Accuracy</i>	48.6	55.13	78.9	92.85	92.72	71.61	49.84	66.46	89.42	77.36	85.75	82.95	71.67	75.4	87.8	92.06	78.02	89.45	86.74	
	<i>Sensitivity</i>	3111.41	1683.61	494.13	80.06	312.69	569.94	1030.24	1508.34	224.58	1291.97	383.94	706.65	1478.31	402.11	196.24	246.97	914.54	391.3	410.29	
	<i>Specificity</i>	0.08	0.22	0.2	0.34	0.17	0.2	0.14	0.19	0.11	0.08	0.17	0.18	0.12	0.14	0.19	0.84	0.34	0.36	0.33	
	<i>Precision</i>	11.85	46.99	5.35	9.74	23.91	5.14	9.75	9.87	68.01	10.94	9.15	46.02	28.22	27.05	4.78	14.32	36.06	17.24	12.84	
	<i>F1-score</i>	96	97.06	96.69	99.48	91.53	97.03	92.81	97.32	96.46	98.23	100	82.1	96.85	95.1	98.11	96.55	98.18	98.3	93.06	
	<i>AUC</i>	6.1	4.18	2.81	4.95	17.76	3.03	4.98	7.33	27.6	2.8	2.62	26.84	13.41	7.16	3.57	6.79	4.08	3.7	10.97	
	<i>FP/h</i>	6.53	46.41	2.7	5.29	9.19	2.25	5.31	2.96	65.42	8.6	6.87	37.36	19.44	23.04	1.3	8.66	34.51	14.56	2.38	
	<i>Prediction Time</i>	12.22	62.79	5.25	10.05	16.71	4.39	10.05													

Table S9: TUH (Generalised) validation performance before post-processing.

Classifier	Metric	P7	P9	P10	P14	P19	P35	P39	P40	P47	P56	P61	P62	P65
LightGBM	<i>Accuracy</i>	93.04	91.7	92.58	80.51	79.83	80.9	88.63	97.17	60.59	85.01	97.84	51.54	67.4
	<i>Sensitivity</i>	86.96	36	65.24	78.02	65.2	46.3	89.22	34.32	46.86	73.29	53.92	97.83	32.14
	<i>Specificity</i>	93.25	94.69	98.19	80.71	98.48	81.28	86.39	99.76	75.01	86.83	98.92	47.81	89.67
	<i>Precision</i>	30.3	26.73	88.07	23.78	98.2	2.65	96.16	85.26	66.31	46.27	55.08	13.15	66.25
	<i>F1-score</i>	44.94	30.68	74.96	36.46	78.37	5.01	92.56	48.94	54.92	56.73	54.49	23.18	43.28
	<i>AUC</i>	90.1	65.35	81.72	79.36	81.84	63.79	87.81	67.04	60.93	80.06	76.42	72.82	60.9
	<i>FP/h</i>	235.23	181.35	54.15	644.84	24.1	666.61	101.69	8.45	438.9	410.66	38.04	1738.59	228.07
	<i>Prediction Time</i>	0.02	0.02	0.02	0.02	0.03	0.02	0.01	0.03	0.09	0.02	0.05	0.02	0.02
MLP	<i>Accuracy</i>	79.19	71.52	79.98	56.36	73.63	22.97	88.06	94.58	56.08	47.54	96.43	19.74	73.13
	<i>Sensitivity</i>	93.48	66.22	82.32	88.79	64.53	90.74	98.57	58.9	71.76	95.03	71.39	91.3	91.36
	<i>Specificity</i>	78.71	71.81	79.5	53.86	85.22	22.22	47.96	96.05	39.61	40.19	97.05	13.96	61.62
	<i>Precision</i>	12.91	11.22	45.15	12.93	84.76	1.27	87.85	38.08	55.51	19.74	37.32	7.89	60.04
	<i>F1-score</i>	22.69	19.19	58.32	22.58	73.28	2.5	92.9	46.26	62.6	32.69	49.02	14.53	72.46
	<i>AUC</i>	86.09	69.01	80.91	71.33	74.88	56.48	73.27	77.48	55.69	67.61	84.22	52.63	76.49
	<i>FP/h</i>	741.48	963.1	612.45	1542.06	234	2769.45	388.98	136.44	1060.46	1864.45	103.71	2865.96	847.01
	<i>Prediction Time</i>	0.12	0.2	0.1	0.17	0.32	0.23	0.08	0.27	0.88	0.08	0.63	0.13	0.2
RNN	<i>Accuracy</i>	6.38	8.26	24.99	12.71	55.37	4.07	79.99	7.65	51.28	25.32	10.2	18.46	39.01
	<i>Sensitivity</i>	100	98.67	98.48	98.92	95.61	92.59	97.26	98.31	96.17	92.45	97.59	98.88	98.72
	<i>Specificity</i>	3.2	3.4	9.93	5.9	4.87	3.1	14.24	3.91	4.05	14.79	8.05	11.99	2.14
	<i>Precision</i>	3.38	5.21	18.3	7.67	55.78	1.04	81.19	4.05	51.33	14.54	2.54	8.29	38.38
	<i>F1-score</i>	6.55	9.9	30.86	14.24	70.45	2.06	88.5	7.77	66.94	25.13	4.96	15.29	55.27
	<i>AUC</i>	51.6	51.03	54.2	52.41	50.24	47.84	55.75	51.11	50.11	53.62	52.82	55.44	50.43
	<i>FP/h</i>	3370.49	3300.2	2691.14	3139.49	1518.63	3450.57	642.49	3322.29	1683.2	2651.66	3230.66	2932.39	2178.22
	<i>Prediction Time</i>	0.24	0.9	0.41	0.6	1.38	1.02	0.34	1.22	4.15	0.26	2.85	0.51	0.9
CNN1D	<i>Accuracy</i>	89.44	91.38	85.33	89.11	49.31	44.41	91.11	51.53	62.45	82.18	65.18	59.71	52.7
	<i>Sensitivity</i>	72.22	31.11	78.66	69.19	11.43	68.52	96.72	78.39	84.29	9.43	96.69	73.74	70.19
	<i>Specificity</i>	90.02	94.62	86.7	90.68	96.85	44.14	69.79	50.43	39.48	93.59	64.4	58.58	41.91
	<i>Precision</i>	19.7	23.73	54.78	36.99	82.01	1.33	92.41	6.12	59.44	18.75	6.26	12.52	42.72
	<i>F1-score</i>	30.95	26.92	64.58	48.21	20.06	2.61	94.52	11.35	69.71	12.55	11.75	21.41	53.12
	<i>AUC</i>	81.12	62.87	82.68	79.94	54.14	56.33	83.25	64.41	61.88	51.51	80.54	66.16	56.05
	<i>FP/h</i>	347.54	183.8	397.51	310.81	50.23	1988.96	226.3	1713.97	1061.67	199.49	1250.81	1380.12	1293.06
	<i>Prediction Time</i>	4.25	14.1	6.2	8.1	21.29	16.2	4.45	19.13	64.96	3.78	44.21	7.71	14.34

Note. All metrics are in percent apart from FP/h and prediction time.

Table S10: TUH (Generalised) test performance before post-processing.

Classifier	Metric	P3	P17	P20	P23	P26	P31	P32	P33	P38	P50	P51	P54	P64
LightGBM	<i>Accuracy</i>	81.58	83.8	97.6	95.21	57.4	90.88	95.39	95.09	82.92	80.93	94.35	93.38	47.75
	<i>Sensitivity</i>	86.86	38.82	82.82	100	54.25	57.34	93.62	5.59	0.2	1.47	91.46	7.1	0.36
	<i>Specificity</i>	80.75	91.77	98.46	95.16	61.68	99.74	95.45	98.88	96.37	90.12	94.48	97.6	98.06
	<i>Precision</i>	41.27	45.54	75.81	15.04	65.76	98.35	43.14	17.38	0.88	1.7	43.1	12.62	16.44
	<i>F1-score</i>	55.96	41.91	79.16	26.14	59.46	72.44	59.06	8.46	0.33	1.58	58.59	9.09	0.7
	<i>AUC</i>	83.81	65.29	90.64	97.58	57.96	78.54	94.53	52.23	48.28	45.8	92.97	52.35	49.21
	<i>FP/h</i>	599.57	251.59	52.47	172.59	585.39	7.26	157.94	38.82	112.5	318.76	189.88	82.44	33.92
	<i>Prediction Time</i>	0.02	0.1	0.02	0.02	0.02	0.03	0.01	0.2	0.02	0.03	0.02	0.02	0.04
MLP	<i>Accuracy</i>	70.09	76.76	66.92	56.94	62.19	87.52	39.18	82.28	83.73	82.08	71.44	59.54	48.05
	<i>Sensitivity</i>	95.71	48.39	91.19	95	81.48	71.74	100	31.7	2.4	58.26	95.12	82.51	7.93
	<i>Specificity</i>	66.1	81.78	65.5	56.61	36.02	91.69	36.94	84.42	96.95	84.83	70.36	58.42	90.64
	<i>Precision</i>	30.54	32.02	13.36	1.84	63.34	69.54	5.52	7.92	11.32	30.76	12.79	8.84	47.36
	<i>F1-score</i>	46.3	38.54	23.31	3.61	71.27	70.62	10.47	12.67	3.95	40.27	22.54	15.96	13.59
	<i>AUC</i>	80.9	65.09	78.35	75.81	58.75	81.72	68.47	58.06	49.67	71.55	82.74	70.47	49.29
	<i>FP/h</i>	1056.07	557.04	1173.48	1548.75	977.33	236.58	2189.41	538.11	94.42	489.43	1020.35	1427.18	163.48
	<i>Prediction Time</i>	0.14	0.98	0.19	0.13	0.17	0.32	0.08	2.16	0.16	0.29	0.1	0.19	0.58
RNN	<i>Accuracy</i>	19.07	18.99	18.41	4.98	57.25	29.61	7.43	6.41	20.3	16.07	10.04	11.98	50.29
	<i>Sensitivity</i>	98.64	96.52	96.41	95	97.56	97.59	97.78	96.91	93.61	97.05	100	96.72	91.99
	<i>Specificity</i>	6.54	5.32	14	4.23	2.5	11.64	4.17	2.59	8.43	6.71	5.93	7.84	6.32
	<i>Precision</i>	14.26	15.24	5.95	0.82	57.61	22.59	3.55	4.03	14.2	10.74	4.63	4.87	50.87
	<i>F1-score</i>	24.91	26.33	11.21	1.63	72.44	36.69	6.85	7.74	24.66	19.34	8.84	9.28	65.51
	<i>AUC</i>	52.59	50.92	55.21	49.61	50.03	54.61	50.97	49.75	51.02	51.88	52.97	52.28	49.16
	<i>FP/h</i>	2906.67	2897.49	2930.48	3419.25	1488.48	2516.01	3329.72	3364.63	2837.15	3010.37	3238.63	3163.16	1641.65
	<i>Prediction Time</i>	0.51	4.5	0.76	0.51	0.73	1.41	0.28	10.13	0.75	1.41	0.39	0.89	2.72
CNN1D	<i>Accuracy</i>	61.3	58.79	72.53	57.97	55.47	87.16	88.85	27.73	61.26	66.54	78.01	41.62	57.28
	<i>Sensitivity</i>	68.37	57.1	70.26	95	75.82	74.22	97.78	82.34	19.16	58.85	93.75	59.56	67.7
	<i>Specificity</i>	60.18	59.09	72.66	57.66	27.84	90.58	88.53	25.42	68.08	67.43	77.3	40.74	46.31
	<i>Precision</i>	21.29	19.76	12.66	1.84	58.8	67.57	23.53	4.46	8.86	17.29	15.86	4.68	57.07
	<i>F1-score</i>	32.47	29.36	21.46	3.62	66.23	70.74	37.93	8.46	12.11	26.72	27.12	8.67	61.93
	<i>AUC</i>	64.27	58.1	71.46	76.33	51.83	82.4	93.16	53.88	43.62	63.14	85.52	50.15	57
	<i>FP/h</i>	1238.33	1251.95	931.8	1511.7	1101.66	268.15	398.45	2575.89	989.1	1050.99	781.67	2033.99	940.94
	<i>Prediction Time</i>	6.92	70.56	11.67	7.77	11.46	22.17	4.16	160.68	12.32	20.9	5.88	12.55	41.51

Note. All metrics are in percent apart from FP/h and prediction time.

Table S11: TUH (Generalised) training performance after post-processing.

Classifier	Metric	P1	P2	P4	P5	P6	P8	P11	P12	P13	P15	P16	P18	P21	P22	P24	P25	P27	P28	P29	P30
LightGBM	Accuracy	96.49	96.94	93.67	83.24	87.44	80.81	29.63	93.18	93.6	80.73	88.3	91.45	96.38	97.21	75.35	98.21	92.32	99.26	97.22	94.42
	Sensitivity	88.42	96.43	86.31	38.4	6.68	0	29.63	66.84	38.42	71.47	51.39	89.04	93.1	91.4	37.87	96.61	55.31	62.86	55.93	7.39
	Specificity	98.2	96.94	100	100	99.81	95.37	-	99.91	100	86.32	91.68	94.14	96.44	100	100	98.28	99.17	100	99.5	100
	Precision	91.25	18.62	100	100	84.44	0	100	99.48	100	75.9	36.13	94.43	29.67	100	100	71.25	92.5	100	86.09	100
	F1-score	89.81	31.21	92.65	55.49	12.38	0	45.71	79.96	55.51	73.62	42.43	91.66	45	95.51	54.94	82.01	69.22	77.19	67.81	13.76
	AUC	93.31	96.68	93.15	69.2	53.24	47.69	-	83.38	69.21	78.89	71.54	91.59	94.77	95.7	68.93	97.45	77.24	81.43	77.72	53.69
	FP/h	53.36	109.32	0	0	5.88	141.08	0	2.57	0	307.32	274.33	99.71	126.25	0	0	59.27	25.22	0	17.01	0
	Prediction Time	0.02	0.02	0.01	0.01	0.02	0.01	0.01	0.03	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.02	0.02	0.04	0.03
	MLP	Accuracy	91.9	99.9	89.36	93.55	87.37	83.45	40.74	56.09	95.69	84.04	95.76	90.78	95.84	98.95	89.51	98.5	91.36	99.07	97.63
Sensitivity		88.42	96.43	76.97	85.82	6.15	0	40.74	29.82	58.58	75.31	49.4	88.86	0	97.31	73.57	96.61	76.53	53.33	60.45	54.26
Specificity		92.63	99.92	100	96.44	99.81	98.48	-	62.81	100	89.3	100	92.93	97.38	99.74	100	98.58	94.11	100	99.68	98.36
Precision		71.79	90	100	90	83.33	0	100	17.02	100	80.94	100	93.34	0	99.45	100	75	70.62	100	91.19	67.97
F1-score		79.24	93.1	86.99	87.86	11.46	0	57.89	21.67	73.88	78.02	66.13	91.04	0	98.37	84.77	84.44	73.46	69.57	72.71	60.35
AUC		90.53	98.18	88.49	91.13	52.98	49.24	-	46.32	79.29	82.31	74.7	90.89	48.69	98.53	86.79	97.6	85.32	76.67	80.06	76.31
FP/h		218.77	2.78	0	93.41	5.88	46.22	0	1066.24	0	240.21	0	120.34	92.71	6.28	0	48.96	179.03	0	10.99	55.42
Prediction Time		0.11	0.18	0.07	0.08	0.2	0.09	0.03	0.18	0.17	0.25	0.14	0.11	0.1	0.05	0.08	0.08	0.19	0.23	0.44	0.25
RNN		Accuracy	56.09	58.59	45.38	72.79	43.92	71.22	100	23.44	61.32	73.23	45.65	70.74	86.21	55.67	75.58	63.72	58.46	65.56	42.81
	Sensitivity	72.18	100	74.67	0	72.41	44.2	100	77.8	82.02	67.38	43.65	83.36	0	89.78	63.43	100	70.2	83.81	87.52	67.9
	Specificity	52.53	58.28	20.18	100	39.55	76.23	-	9.49	58.92	76.74	45.83	56.55	87.55	39.28	83.55	62.11	56.28	65.19	40.35	33.55
	Precision	25.15	1.72	44.59	0	15.51	25.64	100	18.07	18.81	63.48	6.92	68.33	0	41.54	71.69	10.47	22.98	4.66	7.48	6.14
	F1-score	37.31	3.39	55.83	0	25.54	32.45	100	29.33	30.6	65.37	11.94	75.1	0	56.8	67.31	18.95	34.63	8.82	13.79	11.27
	AUC	62.36	79.14	47.43	50	55.98	60.21	-	43.65	70.47	72.06	44.74	69.96	43.77	64.53	73.49	81.05	63.24	74.5	63.94	50.72
	FP/h	1399.62	1490.89	1544.76	0	1887	721.87	0	2592.86	1325.13	523.29	1785.34	736.2	441.44	1476.44	357.41	1306.22	1327.17	1228.4	2035.2	2248.31
	Prediction Time	0.34	0.64	0.18	0.31	0.88	0.27	0.04	0.47	0.76	1.46	0.55	0.44	0.41	0.13	0.28	0.3	0.86	1.07	2.12	1.19
	CNN1D	Accuracy	85.02	99.27	53.76	72.79	89.73	84.36	0	42.86	93.06	72.79	94.9	78.96	97.2	83.94	65.01	96.49	89.76	99.68	96.6
Sensitivity		19.01	0	0	0	27.77	0	0	56.5	33.24	40.71	39.59	65.62	89.66	50.54	11.72	34.48	36.16	84.76	39.7	6.82
Specificity		99.61	100	100	100	99.22	100	-	39.36	100	92.03	100	93.97	97.32	100	100	99.24	99.72	99.98	99.74	97.36
Precision		91.53	0	0	0	84.49	0	0	19.3	100	75.4	100	92.45	34.21	100	100	66.67	95.98	98.89	89.36	14.2
F1-score		31.49	0	0	0	41.8	0	0	28.77	49.9	52.88	56.73	76.76	49.52	67.14	20.98	45.45	52.53	91.28	54.97	9.21
AUC		59.31	50	50	50	63.49	50	-	47.93	66.62	66.37	69.8	79.8	93.49	75.27	55.86	66.86	67.94	92.37	69.72	52.09
FP/h		11.47	0	0	0	24.38	0	0	1737.36	0	179.36	0	102.15	95.14	0	0	26.34	8.53	0.68	8.89	89.29
Prediction Time		5.3	10.27	3.37	5.18	13.67	3.72	0.81	9.16	11.31	19.33	8.28	7.34	6.07	1.84	4.06	5.03	14.22	16.88	33.26	19.33
LightGBM		Classifier	Metric	P34	P36	P37	P41	P42	P43	P44	P45	P46	P48	P49	P52	P53	P55	P57	P58	P59	P60
	Accuracy	93.67	98.96	96.39	98.95	91.65	96.64	92.52	94.91	89.65	94.25	97.07	91.95	89.98	78.15	98.7	95.34	92.94	96.15	99.1	
	Sensitivity	0	97.88	0	81.56	4.41	49.5	0	47.86	88.27	64.66	75	84.8	58.5	2.82	0	54.33	85.66	75.65	74.29	
	Specificity	100	99.89	99.08	99.88	99.63	97.72	97.83	96.3	91.61	97.02	98.66	95.76	96.85	100	100	99.44	96.72	99.57	99.69	
	Precision	0	99.86	0	97.21	52	33.33	0	27.72	93.73	66.96	80.08	91.41	80.24	100	0	90.62	93.11	96.69	84.97	
	F1-score	0	98.86	0	88.7	8.12	39.84	0	35.11	90.92	65.79	77.46	87.98	67.67	5.49	0	67.93	89.23	84.88	79.27	
	AUC	50	98.88	49.54	90.72	52.02	73.61	48.92	72.08	89.94	80.84	86.83	90.28	77.68	51.41	50	76.88	91.19	87.61	86.99	
	FP/h	0	2.2	32.24	4.26	12.26	80.16	73.76	129.24	124.77	98.16	45	99.63	92.98	0	0	18.41	77.85	13.33	10.92	
	Prediction Time	0.01	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.09	0.03	0.03	
MLP	Accuracy	43.88	64.8	97.51	98.89	96.76	96.73	94.57	95.75	65.2	95.14	97.47	78.74	86	77.87	99.76	97.08	90.33	97.93	98.76	
	Sensitivity	84.21	90.51	33.06	79.48	70.51	0	0	30.77	44.88	81.9	67.06	74.2	72.05	16.14	81.13	83.07	83.76	85.68	67.43	
	Specificity	41.16	42.86	99.31	99.93	99.16	98.95	100	97.67	94.07	96.37	99.66	81.16	89.05	95.77	100	98.48	93.73	99.97	99.5	
	Precision	8.82	57.49	57.14	98.39	88.51	0	0	28.12	91.49	67.86	93.37	67.69	58.96	52.55	100	84.51	87.39	99.81	76.13	
	F1-score	15.96	70.31	41.88	87.93	78.49	0	0	29.39	60.22	74.22	78.06	70.8	64.85	24.7	89.58	83.78	85.54	92.21	71.52	
	AUC	62.68	66.68	66.18	89.71	84.84	49.48	50	64.22	69.48	89.14	83.36	77.68	80.55	55.96	90.57	90.77	88.75	92.83	83.46	
	FP/h	1984.35	1109.72	24.18	2.37	27.59	36.87	0	81.44	88.17	119.38	11.49	442.68	323.55	117.97	0	49.83	148.55	0.83	17.57	
	Prediction Time	0.08	0.22	0.2	0.34	0.17	0.2	0.14	0.19	0.11	0.08	0.17	0.18	0.12	0.14	0.19	0.84	0.34	0.36	0.33	
	RNN	Accuracy	54.05	53.43	17.94	40.32	79.71	25.18	37.55	51.52	45.75	28.13	22.02	65.3	77.63	47.95	26.27	47.17	49.93	35.84	57.09
Sensitivity		70.67	78.92	72.73	93.77	37.63	94.06	64.07	58.04	14	92.04	100	0	67.05	75.2	90.57	80.27	87.3	89.24	47.98	
Specificity		52.91	31.63	16.41	37.47	83.53	23.6	36.03	51.33	90.84	22.18	16.42	100	79.91	39.98	25.43	44.14	30.69	26.92	57.3	
Precision		9.3	49.67	2.37	7.41	17.21	2.75	5.44	3.34	68.46	9.92	7.91	0	41.86	26.8	1.55	11.63	39.33	16.94	2.56	
F1-score		16.43	60.97	4.58	13.73	23.62	5.35	10.02	6.32	23.25	17.92	14.67	0	51.54	39.52	3.05	20.31	54.23	28.47	4.86	
AUC		61.79	55.28	44.57	65.62	60.58	58.83	50.05	54.68	52.42	57.11	58.21	50	73.48	57.59	58	62.21	58.99	58.08	52.64	
FP/h		1586.7	1326.83	2927.72	2137.07	543.36	2688.59	2178.02	1702.59	136.29	2562.9</										

Table S12: TUH (Generalised) validation performance after post-processing.

Classifier	Metric	P7	P9	P10	P14	P19	P35	P39	P40	P47	P56	P61	P62	P65
LightGBM	<i>Accuracy</i>	97.37	95.37	82.99	96.23	78.02	97.76	78.39	96.91	55.39	93.26	98.49	84.48	63.12
	<i>Sensitivity</i>	84.78	9.33	0	71.98	60.76	27.78	72.73	22.03	14.76	57.76	37.35	89.13	4.69
	<i>Specificity</i>	97.8	100	100	98.1	100	98.53	100	100	98.04	98.75	100	84.11	100
	<i>Precision</i>	56.52	100	0	74.55	100	17.24	100	100	88.77	87.74	100	31.18	100
	<i>F1-score</i>	67.83	17.07	0	73.25	75.59	21.28	84.21	36.11	25.31	69.66	54.39	46.2	8.95
	<i>AUC</i>	91.29	54.67	50	85.04	80.38	63.16	86.36	61.02	56.4	78.26	68.67	86.62	52.34
	<i>FP/h</i>	76.7	0	0	63.37	0	52.23	0	0	34.41	38.97	0	529.33	0
<i>Prediction Time</i>	0.02	0.02	0.02	0.02	0.03	0.02	0.01	0.03	0.09	0.02	0.05	0.02	0.02	
MLP	<i>Accuracy</i>	97.59	95.92	82.99	92.65	57.34	74.17	95.34	96.7	58.87	94.59	98.57	40.58	86.65
	<i>Sensitivity</i>	93.48	25.33	0	61.21	23.87	83.33	94.21	17.8	28.35	86.34	40.36	63.04	65.55
	<i>Specificity</i>	97.72	99.71	100	95.08	100	74.07	99.66	99.95	90.92	95.87	100	38.76	99.96
	<i>Precision</i>	58.11	82.61	0	48.97	100	3.41	99.91	93.33	76.62	76.37	100	7.68	99.91
	<i>F1-score</i>	71.67	38.78	0	54.41	38.53	6.56	96.97	29.89	41.39	81.05	57.51	13.69	79.16
	<i>AUC</i>	95.6	62.52	50	78.14	61.93	78.7	96.93	58.87	59.63	91.1	70.18	50.9	82.76
	<i>FP/h</i>	79.26	9.8	0	164.55	0	923.39	2.54	1.81	159.54	128.89	0	2039.81	0.85
<i>Prediction Time</i>	0.12	0.2	0.1	0.17	0.32	0.23	0.08	0.27	0.88	0.08	0.63	0.13	0.2	
RNN	<i>Accuracy</i>	26.96	23.94	57.39	43.52	55.78	23.35	82.8	33.07	51.17	88.24	54.67	75.3	43.77
	<i>Sensitivity</i>	100	85.78	57.32	90.27	66.28	61.11	81.2	88.56	71.07	41.51	91.27	58.1	94.46
	<i>Specificity</i>	24.48	20.61	57.4	39.83	42.59	22.94	88.89	30.78	30.22	95.56	53.77	76.68	12.47
	<i>Precision</i>	4.3	5.49	21.61	10.6	59.17	0.86	96.53	5.01	51.73	59.46	4.63	16.69	39.98
	<i>F1-score</i>	8.24	10.33	31.39	18.98	62.52	1.71	88.21	9.48	59.88	48.89	8.81	25.94	56.19
	<i>AUC</i>	62.24	53.19	57.36	65.05	54.44	42.02	85.05	59.67	50.64	68.54	72.52	67.39	53.47
	<i>FP/h</i>	2629.51	2712.05	1272.78	2007.45	916.47	2744.07	83.24	2393.16	1224.13	138.11	1624.3	776.88	1948.36
<i>Prediction Time</i>	0.24	0.9	0.41	0.6	1.38	1.02	0.34	1.22	4.15	0.26	2.85	0.51	0.9	
CNN1D	<i>Accuracy</i>	96.72	94.89	83	95.88	44.34	98.45	89.02	96.31	62.08	86.45	93.24	92.56	69.82
	<i>Sensitivity</i>	0	0	0	45.95	0	0	86.13	8.9	31.04	0	89.76	0	24.15
	<i>Specificity</i>	100	100	100	99.83	100	99.53	100	99.91	94.74	100	93.33	100	98.01
	<i>Precision</i>	0	0	0	95.51	0	0	100	80.77	86.14	0	24.85	0	88.22
	<i>F1-score</i>	0	0	0	62.04	0	0	92.55	16.03	45.63	0	38.93	0	37.92
	<i>AUC</i>	50	50	50	72.89	50	49.77	93.07	54.41	62.89	50	91.54	50	61.08
	<i>FP/h</i>	0	0	0	5.7	0	16.68	0	3.02	92.2	0	234.35	0	44.3
<i>Prediction Time</i>	4.25	14.1	6.2	8.1	21.29	16.2	4.45	19.13	64.96	3.78	44.21	7.71	14.34	

Note. All metrics are in percent apart from FP/h and prediction time.

Table S13: TUH (Generalised) test performance after post-processing.

Classifier	Metric	P3	P17	P20	P23	P26	P31	P32	P33	P38	P50	P51	P54	P64
LightGBM	<i>Accuracy</i>	92.38	87.06	98.59	99.28	44.22	86.45	96.29	95.89	86.02	89.63	98.88	95.34	48.51
	<i>Sensitivity</i>	72.12	26.78	74.45	100	18.77	35.29	93.62	0.59	0	0	90.24	0	0
	<i>Specificity</i>	95.53	97.74	100	99.27	78.76	99.98	96.39	99.92	100	100	99.28	100	100
	<i>Precision</i>	71.54	67.78	100	54.05	54.52	99.81	48.89	23.08	0	0	85.06	0	0
	<i>F1-score</i>	71.83	38.39	85.35	70.18	27.92	52.14	64.23	1.15	0	0	87.57	0	0
	<i>AUC</i>	83.83	62.26	87.22	99.64	48.76	67.63	95	50.25	50	50	94.76	50	50
	<i>FP/h</i>	139.16	69	0	25.97	324.43	0.52	125.26	2.86	0	0	24.93	0	0
<i>Prediction Time</i>	0.02	0.1	0.02	0.02	0.02	0.03	0.01	0.2	0.02	0.03	0.02	0.02	0.02	0.04
MLP	<i>Accuracy</i>	92.85	87.37	98.52	97.88	48.53	86.99	94.86	96.09	86.02	90.23	96.64	88.8	48.46
	<i>Sensitivity</i>	94.91	17.68	86.78	95	34.47	38.04	100	8.98	0	5.75	86.59	40.44	1.14
	<i>Specificity</i>	92.53	99.72	99.2	97.9	67.61	99.93	94.67	99.77	100	100	97.1	91.17	98.68
	<i>Precision</i>	66.42	91.74	86.4	27.94	59.08	99.28	40.87	62.46	0	100	57.72	18.27	47.8
	<i>F1-score</i>	78.15	29.65	86.59	43.18	43.54	55.01	58.02	15.7	0	10.88	69.27	25.17	2.23
	<i>AUC</i>	93.72	58.7	92.99	96.45	51.04	68.98	97.33	54.38	50	52.88	91.84	65.8	49.91
	<i>FP/h</i>	232.8	8.62	27.11	74.84	494.71	2.08	185.17	7.88	0	0	99.73	303.21	23.08
<i>Prediction Time</i>	0.14	0.98	0.19	0.13	0.17	0.32	0.08	2.16	0.16	0.29	0.1	0.19	0.58	
RNN	<i>Accuracy</i>	53.98	44.64	78.64	29.96	54.67	79.8	31.27	20.45	49.33	57.88	40.7	54.18	48.76
	<i>Sensitivity</i>	97.96	75.36	75.38	95	84.43	83.25	97.78	81.94	47.7	79.2	100	78.14	47.44
	<i>Specificity</i>	47.05	39.22	78.82	29.41	14.26	78.89	28.87	17.85	49.6	55.42	37.99	53.01	50.15
	<i>Precision</i>	22.57	17.95	16.72	1.11	57.22	51.04	4.73	4.04	13.29	17.05	6.86	7.51	50.08
	<i>F1-score</i>	36.69	28.99	27.37	2.2	68.21	63.28	9.02	7.71	20.78	28.06	12.83	13.7	48.72
	<i>AUC</i>	72.51	57.29	77.1	62.21	49.34	81.07	63.32	49.9	48.65	67.31	69	65.58	48.79
	<i>FP/h</i>	1646.67	1860.13	721.77	2520	1309	601.12	2471.52	2837.47	1561.74	1438.57	2134.86	1612.72	873.61
<i>Prediction Time</i>	0.51	4.5	0.76	0.51	0.73	1.41	0.28	10.13	0.75	1.41	0.39	0.89	2.72	
CNN1D	<i>Accuracy</i>	88.61	88.17	97.62	99.63	50.72	86.44	96.36	75.81	86.07	90.37	99.62	95.34	58.65
	<i>Sensitivity</i>	26.87	26.09	55.38	90	34.48	35.15	93.33	38.86	0	7.23	91.25	0	25.07
	<i>Specificity</i>	98.34	99.13	100	99.71	72.77	100	96.47	77.37	100	99.98	100	100	94.04
	<i>Precision</i>	71.82	84.03	100	72	63.23	100	48.84	6.76	0	98	100	0	81.6
	<i>F1-score</i>	39.11	39.82	71.29	80	44.62	52.01	64.12	11.52	0	13.46	95.42	0	38.36
	<i>AUC</i>	62.6	62.61	77.69	94.85	53.62	67.57	94.9	58.11	50	53.61	95.62	50	59.56
	<i>FP/h</i>	51.67	26.76	0	10.46	415.7	0	122.6	781.8	0	0.55	0	0	104.46
<i>Prediction Time</i>	6.92	70.56	11.67	7.77	11.46	22.17	4.16	160.68	12.32	20.9	5.88	12.55	41.51	

Note. All metrics are in percent apart from FP/h and prediction time.