

1. Supplements. Evaluation metrics for the models

Table S.1. Evaluation metrics and results of all models.

Model	Class	TP	FP	FN	TN	Support	Precision	Recall	F1-score	Accuracy	Specificity	Error rate	Type I error	Type II error	
KNN	Cat. 1	34	7	1	2780	35	0.83	0.97	0.89	1.00	1.00	0.00	0.00	0.03	
	Cat. 2	28	1	19	2774	47	0.97	0.60	0.74	0.99	1.00	0.01	0.00	0.40	
	Cat. 3	911	30	29	1852	940	0.97	0.97	0.97	0.98	0.98	0.02	0.02	0.03	
	Cat. 4	942	30	40	1810	982	0.97	0.96	0.96	0.98	0.98	0.02	0.02	0.04	
	Cat. 5	811	28	7	1976	818	0.97	0.99	0.98	0.99	0.99	0.01	0.01	0.01	
	Total (macro-average)						2822	0.94	0.90	0.92	0.99	0.99	0.01	0.01	0.10
	Total (micro-average)							0.97	0.97	0.97		0.99		0.01	0.03
SGD	Cat. 1	27	0	8	2787	35	1.00	0.77	0.87	1.00	1.00	0.00	0.00	0.23	
	Cat. 2	0	0	47	2775	47	0.00	0.00	0.00	0.98	1.00	0.02	0.00	1.00	
	Cat. 3	626	86	314	1796	940	0.88	0.67	0.76	0.86	0.95	0.14	0.05	0.33	
	Cat. 4	947	418	35	1422	982	0.69	0.96	0.81	0.84	0.77	0.16	0.23	0.04	
	Cat. 5	679	39	139	1965	818	0.95	0.83	0.88	0.94	0.98	0.06	0.02	0.17	
	Total (macro-average)						2822	0.70	0.65	0.67	0.92	0.94	0.08	0.06	0.35
	Total (micro-average)							0.81	0.81	0.81		0.95		0.05	0.19
QDA	Cat. 1	0	0	35	2787	35	0.00	0.00	0.00	0.99	1.00	0.01	0.00	1.00	
	Cat. 2	0	0	47	2775	47	0.00	0.00	0.00	0.98	1.00	0.02	0.00	1.00	
	Cat. 3	0	0	940	1882	940	0.00	0.00	0.00	0.67	1.00	0.33	0.00	1.00	
	Cat. 4	982	1840	0	0	982	0.35	1.00	0.52	0.35	0.00	0.65	1.00	0.00	
	Cat. 5	0	0	818	2004	818	0.00	0.00	0.00	0.71	1.00	0.29	0.00	1.00	
	Total (macro-average)						2822	0.70	0.20	0.10	0.74	0.80	0.26	0.20	0.80
	Total (micro-average)							0.35	0.35	0.35		0.84		0.16	0.65
AdaBoost	Cat. 1	25	293	10	2394	35	0.06	0.71	0.11	0.86	0.86	0.14	0.14	0.29	
	Cat. 2	7	10	40	2765	47	0.41	0.15	0.22	0.98	1.00	0.02	0.00	0.85	
	Cat. 3	85	18	855	1864	940	0.83	0.09	0.16	0.69	0.99	0.31	0.01	0.91	
	Cat. 4	977	13078	5	533	982	0.43	0.99	0.60	0.54	0.29	0.46	0.71	0.01	
	Cat. 5	0	0	818	2004	818	0.00	0.00	0.00	0.71	1.00	0.29	0.00	1.00	
	Total (macro-average)						2822	0.34	0.39	0.37	0.76	0.83	0.24	0.17	0.61
	Total (micro-average)							0.39	0.39	0.39		0.85		0.15	0.61
XGB	Cat. 1	35	0	0	2787	35	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	
	Cat. 2	45	0	2	2775	47	1.00	0.96	0.98	1.00	1.00	0.00	0.00	0.04	
	Cat. 3	931	9	9	1873	940	0.99	0.99	0.99	0.99	1.00	0.01	0.00	0.01	
	Cat. 4	970	8	12	1832	982	0.99	0.99	0.99	0.99	1.00	0.01	0.00	0.01	
	Cat. 5	815	9	3	1995	818	0.99	1.00	0.99	1.00	1.00	0.00	0.00	0.00	

	Total (macro-average)					2822	0.99	0.99	0.99	1.00	1.00	0.00	0.00	0.01
	Total (micro-average)						0.99	0.99	0.99		1.00		0.00	0.01
Logistic regression	Cat. 1	32	2	3	2785	35	0.94	0.91	0.93	1.00	1.00	0.00	0.00	0.09
	Cat. 2	22	1	25	2774	47	0.96	0.47	0.63	0.99	1.00	0.01	0.00	0.53
	Cat. 3	902	51	38	1831	940	0.95	0.96	0.95	0.97	0.97	0.03	0.03	0.04
	Cat. 4	908	81	74	1759	982	0.92	0.92	0.92	0.95	0.96	0.05	0.04	0.08
	Cat. 5	744	79	74	1925	818	0.90	0.91	0.91	0.95	0.96	0.05	0.04	0.09
	Total (macro-average)					2822	0.93	0.84	0.88	0.97	0.98	0.03	0.02	0.16
	Total (micro-average)						0.92	0.92	0.92		0.98		0.02	0.08
SVM	Cat. 1	34	6	1	2781	35	0.85	0.97	0.91	1.00	1.00	0.00	0.00	0.03
	Cat. 2	35	1	12	2774	47	0.97	0.74	0.84	1.00	1.00	0.00	0.00	0.26
	Cat. 3	899	31	41	1851	940	0.97	0.96	0.96	0.97	0.98	0.03	0.02	0.04
	Cat. 4	892	54	90	1786	982	0.94	0.91	0.93	0.95	0.97	0.05	0.03	0.09
	Cat. 5	794	76	24	1928	818	0.91	0.97	0.94	0.96	0.96	0.04	0.04	0.03
	Total (macro-average)					2822	0.93	0.91	0.92	0.98	0.98	0.02	0.02	0.09
	Total (micro-average)						0.94	0.94	0.94		0.99		0.01	0.06
NB	Cat. 1	34	346	1	2441	35	0.09	0.97	0.16	0.88	0.88	0.12	0.12	0.03
	Cat. 2	12	149	35	2626	47	0.07	0.26	0.12	0.93	0.95	0.07	0.05	0.74
	Cat. 3	30	3	910	1879	940	0.91	0.03	0.06	0.68	1.00	0.32	0.00	0.97
	Cat. 4	135	86	847	1754	982	0.61	0.14	0.22	0.67	0.95	0.33	0.05	0.86
	Cat. 5	817	1210	1	794	818	0.40	1.00	0.57	0.57	0.40	0.43	0.60	0.00
	Total (macro-average)					2822	0.42	0.48	0.45	0.75	0.83	0.25	0.17	0.52
	Total (micro-average)						0.36	0.36	0.36		0.84		0.16	0.64
Decision tree	Cat. 1	34	4	1	2783	35	0.89	0.97	0.93	1.00	1.00	0.00	0.00	0.03
	Cat. 2	42	2	5	2773	47	0.95	0.89	0.92	1.00	1.00	0.00	0.00	0.11
	Cat. 3	925	12	15	1870	940	0.99	0.98	0.99	0.99	0.99	0.01	0.01	0.02
	Cat. 4	966	15	16	1825	982	0.98	0.98	0.98	0.99	0.99	0.01	0.01	0.02
	Cat. 5	814	8	4	1996	818	0.99	1.00	0.99	1.00	1.00	0.00	0.00	0.00
	Total (macro-average)					2822	0.96	0.97	0.96	0.99	1.00	0.01	0.00	0.03
	Total (micro-average)						0.99	0.99	0.99		1.00		0.00	0.01
Random forest	Cat. 1	34	3	1	2784	35	0.92	0.97	0.94	1.00	1.00	0.00	0.00	0.03
	Cat. 2	37	1	10	2774	47	0.97	0.79	0.87	1.00	1.00	0.00	0.00	0.21
	Cat. 3	924	17	16	1865	940	0.98	0.98	0.98	0.99	0.99	0.01	0.01	0.02
	Cat. 4	965	14	17	1826	982	0.99	0.98	0.98	0.99	0.99	0.01	0.01	0.02
	Cat. 5	814	13	4	1991	818	0.98	1.00	0.99	0.99	0.99	0.01	0.01	0.00

	Total (macro-average)					2822	0.97	0.94	0.96	0.99	1.00	0.01	0.00	0.06	
	Total (micro-average)						0.98	0.98	0.98		1.00		0.00	0.02	
CatBoost	Cat. 1	31	0	4	2787	35	1.00	0.89	0.94	1.00	1.00	0.00	0.00	0.11	
	Cat. 2	45	4	2	2771	47	0.92	0.96	0.94	1.00	1.00	0.00	0.00	0.04	
	Cat. 3	931	4	9	1878	940	1.00	0.99	0.99	1.00	1.00	0.00	0.00	0.01	
	Cat. 4	974	8	8	1832	982	0.99	0.99	0.99	0.99	1.00	0.01	0.00	0.01	
	Cat. 5	815	10	3	1994	818	0.99	1.00	0.99	1.00	1.00	0.00	0.00	0.00	
	Total (macro-average)						2822	0.98	0.96	0.97	1.00	1.00	0.00	0.00	0.04
	Total (micro-average)							0.99	0.99	0.99		1.00		0.00	0.01

2. Supplements. Formulas for evaluation metrics used in this research

Table S.2. Formulas for evaluation metrics used in research (Sokolova and Lapalme, 2009).

Metric	Calculation	Remark
Precision	Macro-precision = $\frac{1}{K} * \sum_{i=1}^k \left(\frac{TP_i}{TP_i+FP_i} \right)$	Refers to ratio of true positive predictions compared to total positive predictions. Note: K equals the total number of classes.
	Micro-precision = $\frac{\sum_{i=1}^k TP_i}{\sum_{i=1}^k (TP_i+FP_i)}$	
Recall	Macro-recall = $\frac{1}{K} * \sum_{i=1}^k \left(\frac{TP_i}{TP_i+FN_i} \right)$	Refers to the ratio of true positive predictions compared to total actual positives.
	Micro-recall = $\frac{\sum_{i=1}^k TP_i}{\sum_{i=1}^k (TP_i+FN_i)}$	
Accuracy	$\frac{1}{K} * \sum_{i=1}^k \left(\frac{TP_i+TN_i}{TP_i+TN_i+FN_i+FP_i} \right)$	Refers to the number of labels correctly predicted and represents the effectiveness of the classifier.
Modified accuracy	$\frac{1}{K} * \sum_{i=1}^k \left(\frac{TP_i}{TP_i+TN_i+FN_i+FP_i} \right)$	Modified accuracy was calculated by summing up all true positive predictions only and dividing by the total number of instances.
F1-score	$2 * \frac{Precision * Recall}{Precision + Recall}$	Represents the harmonic average of both positive predictive value and true positive rate (precision and sensitivity).
Specificity	Macro-specificity = $\frac{1}{K} * \sum_{i=1}^k \left(\frac{TN_i}{TN_i+FP_i} \right)$	Refers to the true negative rate, in which negative labels are correctly identified.
	Micro-specificity = $\frac{\sum_{i=1}^k TN_i}{\sum_{i=1}^k (TN_i+FP_i)}$	
Type I error	Macro-Type I error = $\frac{1}{K} * \sum_{i=1}^k \left(\frac{FP_i}{TN_i+FP_i} \right)$	Refers to false positive rate, in which negative labels are incorrectly identified.
	Micro-Type I error = $\frac{\sum_{i=1}^k FP_i}{\sum_{i=1}^k (TN_i+FP_i)}$	
Type II error	Macro-Type II error = $\frac{1}{K} * \sum_{i=1}^k \left(\frac{FN_i}{FN_i+TP_i} \right)$	Refers to false positive rate, in which positive labels are incorrectly identified.
	Micro-Type II error = $\frac{\sum_{i=1}^k FN_i}{\sum_{i=1}^k (FN_i+TP_i)}$	
Error rate	$\frac{1}{K} * \sum_{i=1}^k \left(\frac{FP_i+FN_i}{TP_i+TN_i+FN_i+FP_i} \right)$	Average classification error per individual class.