

Supplement Table 4. Prediction performance of identifier according to independent gene set consisting of significant genes selected through statistical analysis.

Features	Classes	Machine learning algorithms												Support samples	
		support vector machine			Decision tree			random forest			Neural network				
		precision	recall	f1-score	precision	recall	f1-score	precision	recall	f1-score	precision	recall	f1-score	Train	Test
Gender	Female	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.97	0.98	1.00	0.93	0.96	2774	1272
	Male	0.99	0.99	0.99	0.99	0.99	0.99	0.97	0.98	0.98	0.93	1.00	0.96	2705	1077
Age	10	0.01	0.5	0.02	0.11	0.12	0.12	0	1	0.01	nan	0.00	0.00	21	9
	20	0.34	0.23	0.27	0.09	0.06	0.07	0	0	0	nan	0.00	0.00	185	76
	30	0.25	0.2	0.22	0.2	0.15	0.17	0.67	0.01	0.02	nan	0.00	0.00	401	185
	40	0.26	0.25	0.26	0.16	0.15	0.16	0	0	0	0.28	0.38	0.32	773	307
	50	0.23	0.26	0.25	0.23	0.26	0.24	0	0	0	0.26	0.44	0.33	1279	549
	60	0.33	0.25	0.29	0.3	0.3	0.3	0.5	0	0.01	0.33	0.53	0.41	1451	644
	70	0.34	0.19	0.25	0.26	0.25	0.25	0	0	0	nan	0.00	0.00	1049	460
	80	0.14	0.03	0.05	0.05	0.07	0.06	0	0	0	nan	0.00	0.00	310	115
	90	0	0	0	0	0	0	0	0	0	nan	0.00	0.00	10	4
Race	NATIVE AMERICAN OR ALASKA NATIVE	0.01	0.2	0.03	0	0	0	0	0	0	nan	0.00	0.00	15	5
	ASIAN	0.93	0.77	0.84	0.32	0.33	0.32	0.83	0.09	0.17	0.82	0.88	0.85	405	194
	BLACK	0.94	0.83	0.88	0.35	0.36	0.35	1	0	0.01	0.68	0.95	0.79	536	244
	NATIVE HAWAIIAN	0	0	0	0	0	0	0	0	0	nan	0.00	0.00	3	4
	WHITE	0.97	0.97	0.97	0.87	0.86	0.87	0.83	0.99	0.9	0.98	0.93	0.96	4520	1902
Cancer type	ACC	0.96	1	0.98	0.69	0.82	0.75	1	0.82	0.9	1.00	0.32	0.48	46	22
	BLCA	0.71	0.99	0.83	0.73	0.74	0.73	0.23	1	0.38	0.64	0.88	0.74	256	127
	BRCA	1	0.99	0.99	0.96	0.95	0.95	1	0.95	0.97	0.94	0.99	0.97	661	297

CEC	0.99	0.91	0.95	0.8	0.67	0.73	1	0.35	0.52	nan	0.00	0.00	163	94
CHOL	0.88	0.64	0.74	0.46	0.55	0.5	0	0	0	nan	0.00	0.00	25	11
COAD	0.86	0.85	0.85	0.75	0.72	0.73	0.82	0.83	0.83	0.73	0.97	0.84	176	78
DLBC	1	1	1	1	0.75	0.86	1	0.65	0.79	nan	0.00	0.00	26	20
ESCA	0.89	0.72	0.8	0.72	0.67	0.7	1	0.46	0.63	nan	0.00	0.00	107	46
GBM	1	1	1	0.98	1	0.99	1	0.91	0.95	0.50	0.02	0.04	94	46
HNSC	1	0.99	0.99	0.9	0.91	0.91	1	0.9	0.95	0.98	0.99	0.98	357	140
KICH	0.95	0.86	0.9	0.71	0.68	0.7	0.9	0.41	0.56	nan	0.00	0.00	42	22
KIRC	0.96	0.95	0.95	0.84	0.86	0.85	0.99	0.9	0.94	0.85	0.97	0.90	371	149
KIRP	0.96	0.92	0.94	0.76	0.75	0.75	0.97	0.8	0.88	0.96	0.92	0.94	199	71
LGG	1	1	1	1	0.98	0.99	0.99	0.98	0.99	0.73	1.00	0.84	366	122
LHC	0.99	0.96	0.97	0.91	0.94	0.92	1	0.91	0.95	0.93	0.97	0.95	250	95
LUAD	0.95	0.97	0.96	0.76	0.81	0.78	0.96	0.71	0.82	0.97	0.73	0.83	284	134
LUSC	0.99	0.87	0.93	0.79	0.73	0.76	1	0.57	0.73	0.51	0.86	0.64	258	121
MESO	1	1	1	0.76	0.76	0.76	1	0.76	0.86	0.71	0.24	0.36	61	21
OV	1	0.99	0.99	0.91	0.93	0.92	1	0.94	0.97	1.00	1.00	1.00	194	80
PAAD	0.98	0.98	0.98	0.73	0.72	0.73	0.97	0.66	0.79	0.98	0.84	0.90	124	50
PCPG	1	0.98	0.99	0.93	0.95	0.94	1	0.98	0.99	1.00	1.00	1.00	130	42
PRAD	1	1	1	1	0.96	0.98	1	0.96	0.98	1.00	0.98	0.99	101	52
READ	0.53	0.43	0.48	0.14	0.17	0.16	0	0	0	nan	0.00	0.00	58	23
SARC	0.96	0.98	0.97	0.88	0.84	0.86	1	0.88	0.94	0.56	0.99	0.71	165	83
SKCM	1	0.94	0.97	0.7	0.58	0.63	1	0.55	0.71	0.33	0.09	0.14	68	33
STAD	0.92	0.95	0.93	0.81	0.81	0.81	0.88	0.88	0.88	0.69	1.00	0.82	251	103
TGCT	1	1	1	0.93	0.84	0.89	1	0.78	0.88	0.48	1.00	0.65	96	32
THCA	1	1	1	0.95	0.98	0.97	1	0.97	0.99	0.98	1.00	0.99	284	120
THYM	1	1	1	0.74	0.83	0.79	1	0.71	0.83	0.95	0.95	0.95	75	42
UCEC	0.91	0.96	0.93	0.68	0.76	0.72	0.93	0.31	0.47	0.81	0.67	0.73	108	45

	UCS	1	0.79	0.88	0.3	0.43	0.35	1	0.07	0.13	1.00	0.21	0.35	42	14
	UVM	0.93	1	0.97	0.75	0.86	0.8	0.93	0.93	0.93	nan	0.00	0.00	41	14
Stage	Stage I	0.47	0.74	0.58	0.49	0.46	0.48	0.37	0.92	0.53	0.65	0.35	0.46	960	416
	Stage II	0.49	0.4	0.44	0.45	0.43	0.44	0.64	0.29	0.4	0.44	0.71	0.54	974	444
	Stage III	0.3	0.23	0.26	0.25	0.31	0.28	0.42	0.04	0.07	0.28	0.27	0.28	775	285
	Stage IV	0.42	0.17	0.24	0.34	0.29	0.31	0.54	0.09	0.15	0.52	0.31	0.39	366	173

