T-Test Results for Gaze-based Prediction of Cognitive Load in Augmented Reality

Table 1: T-tests to test for performance of individual models significantly better than random. Sig(FDR) indicates $q < 0.05$ in either one- or two-tailed test.

Window	Window (ms)	AUC Mean	AUC Std	t-stat	p1	q1	p2	q2	Sig(FDR)
3	50.01	0.6104	0.0993	6.1929	$4.07e-07$	$4.07e-07$	$8.14e-07$	$8.14e-07$	True
6	100.02	0.6428	0.1108	7.1782	$2.74e-08$	$2.99e-08$	$5.49e-08$	$5.98e-08$	True
9	150.03	0.6477	0.1115	7.3766	$1.61e-08$	$1.94e-08$	$3.23e-08$	$3.87e-08$	True
12	200.04	0.6747	0.1100	8.8428	$3.7e-10$	$6.34e-10$	$7.39e-10$	$1.27e-09$	True
15	250.05	0.6709	0.1197	7.9462	$3.6e-09$	$5.35e-09$	$7.2e-09$	$1.07e-08$	True
18	300.06	0.7069	0.1144	10.0664	$1.96e-11$	$4.71e-11$	$3.92e-11$	$9.41e-11$	True
21	350.07	0.6916	0.1087	9.8169	$3.51e-11$	$7.02e-11$	$7.02e-11$	$1.4e-10$	True
24	400.08	0.6993	0.1012	10.9614	$2.59e-12$	$1.03e-11$	$5.17e-12$	$2.07e-11$	True
27	450.09	0.6883	0.1326	7.9045	$4.01e-09$	$5.35e-09$	$8.03e-09$	$1.07e-08$	True
30	500.10	0.7227	0.1094	11.3371	$1.14e-12$	$6.83e-12$	$2.28e-12$	$1.37e-11$	True
33	550.11	0.7347	0.1087	12.0255	$2.65e-13$	$3.18e-12$	$5.29e-13$	$6.35e-12$	True
36	600.12	0.7411	0.1247	10.7620	$4.03e-12$	$1.21e-11$	$8.05e-12$	$2.42e-11$	True

Table 2: Significant Features by Time Window (One-tailed, FDR $q < 0.05$)

Window	**Time (ms)**	**Feature**	***p*-value**	***q*-value**	***t*-stat**
3	50.0	gaze_velocity	$4.44e{-}15$	$4.44e{-}14$	12.18
6	100.0	gaze_velocity	$1.21e{-}16$	$1.21e{-}15$	13.14
9	150.0	gaze_velocity	$2.23e{-}20$	$2.23e{-}19$	13.57
12	200.0	gaze_velocity	$1.27e{-}21$	$1.27e{-}20$	14.42
		dispersion_feature	0.00758	0.0379	2.98
15	250.1	gaze_velocity	$2.03e{-}21$	$2.03e{-}20$	14.28
18	300.1	gaze_velocity	$1.14e{-}21$	$1.14e{-}20$	14.48
		dispersion_feature	0.001779	0.0089	3.48
21	350.1	gaze_velocity	$1.20e{-}22$	$1.20e{-}21$	15.14
		dispersion_feature	0.000213	0.00106	3.96
24	400.1	gaze_velocity	$4.34e{-}26$	$4.34e{-}25$	17.76
		dispersion_feature	$3.66e{-}07$	$1.83e{-}06$	6.23
27	450.1	gaze_velocity	$2.27e{-}25$	$2.27e{-}24$	17.18
		dispersion_feature	$3.15e{-}05$	0.000157	4.65
30	500.1	gaze_velocity	$7.72e{-}29$	$7.72e{-}28$	28.05
		dispersion_feature	$1.12e{-}08$	$6.58e{-}08$	7.53
33	550.1	gaze_velocity	$2.95e{-}27$	$2.95e{-}26$	18.71
		dispersion_feature	$1.43e{-}06$	$7.17e{-}06$	5.74
		fix_dur	0.0115	0.0385	2.39
36	600.1	gaze_velocity	$3.53e{-}29$	$3.53e{-}28$	22.35
		dispersion_feature	$1.31e{-}06$	$6.56e{-}05$	5.77
		fix_dur	$6.01e{-}05$	0.0002	4.42

Table 3: T-tests to test for performance of group models significantly better than random (*** indicates $p < 0.05$)

Window	Window (ms)	AUC Mean	AUC Std	t-statistic	p (1-tailed)	p (2-tailed)	Sig ($\alpha = 0.05$)
3	50.01	0.6684	0.0657	14.2828	0.000000***	0.000000***	Yes
6	100.02	0.6932	0.0675	15.9465	0.000000***	0.000000***	Yes
9	150.03	0.7046	0.0721	15.8825	0.000000***	0.000000***	Yes
12	200.04	0.7202	0.0651	18.8422	0.000000***	0.000000***	Yes
15	250.05	0.7237	0.0687	18.1291	0.000000***	0.000000***	Yes
18	300.06	0.7345	0.0662	19.7182	0.000000***	0.000000***	Yes
21	350.07	0.7397	0.0651	20.4952	0.000000***	0.000000***	Yes
24	400.08	0.7525	0.0624	22.5314	0.000000***	0.000000***	Yes
27	450.09	0.7649	0.0618	23.8581	0.000000***	0.000000***	Yes
30	500.10	0.7856	0.0748	21.2506	0.000000***	0.000000***	Yes
33	550.11	0.7915	0.0674	24.0730	0.000000***	0.000000***	Yes
36	600.12	0.7929	0.0699	23.3318	0.000000***	0.000000***	Yes

Table 4: Comparison: Gaze Velocity Individual Model vs. All Features Model

Window	Gaze Vel	All Feat	Diff	t-stat	p-value	Sig	Outperforms
3	0.6395	0.6104	0.0291	1.2913	$1.01e-01$	ns	YES
6	0.6637	0.6428	0.0209	0.8440	$2.01e-01$	ns	YES
9	0.6837	0.6477	0.0360	1.4861	$7.13e-02$	ns	YES
12	0.7020	0.6747	0.0273	1.1237	$1.33e-01$	ns	YES
15	0.7101	0.6709	0.0393	1.5025	$6.92e-02$	ns	YES
18	0.7231	0.7069	0.0162	0.6284	$2.66e-01$	ns	YES
21	0.7227	0.6916	0.0311	1.2695	$1.04e-01$	ns	YES
24	0.7423	0.6993	0.0430	1.8856	$3.19e-02$	*	YES
27	0.7481	0.6883	0.0598	2.1427	$1.84e-02$	*	YES
30	0.7846	0.7227	0.0619	2.5477	$6.67e-03$	**	YES
33	0.7875	0.7347	0.0529	2.1213	$1.88e-02$	*	YES
36	0.7961	0.7411	0.0550	2.0534	$2.24e-02$	*	YES

Table 5: Unpaired (Independent) t-tests: Individual vs Group modeling at each window size

Window	**ms**	**N1**	**N2**	**Indiv**	**Group**	**Diff**	**95% CI**	**t**	**p-value**	**d**	**Sig**
3	50.0	31	31	0.6104	0.6684	-0.0580	[-0.1008, -0.0153]	-2.714	0.0087	-0.689	**
6	100.0	31	31	0.6428	0.6932	-0.0504	[-0.0970, -0.0038]	-2.165	0.0344	-0.550	*
9	150.0	31	31	0.6477	0.7046	-0.0569	[-0.1046, -0.0092]	-2.385	0.0202	-0.606	*
12	200.0	31	31	0.6747	0.7202	-0.0455	[-0.0914, +0.0004]	-1.982	0.0521	-0.503	
15	250.1	31	31	0.6709	0.7237	-0.0528	[-0.1024, -0.0032]	-2.129	0.0373	-0.541	*
18	300.1	31	31	0.7069	0.7345	-0.0276	[-0.0751, +0.0199]	-1.162	0.2500	-0.295	
21	350.1	31	31	0.6916	0.7397	-0.0481	[-0.0936, -0.0025]	-2.112	0.0388	-0.537	*
24	400.1	31	31	0.6993	0.7525	-0.0532	[-0.0959, -0.0104]	-2.489	0.0156	-0.632	*
27	450.1	31	31	0.6883	0.7649	-0.0766	[-0.1291, -0.0240]	-2.913	0.0050	-0.740	**
30	500.1	31	31	0.7227	0.7856	-0.0629	[-0.1105, -0.0153]	-2.642	0.0105	-0.671	*
33	550.1	31	31	0.7347	0.7915	-0.0569	[-0.1028, -0.0109]	-2.475	0.0162	-0.629	*
36	600.1	31	31	0.7411	0.7929	-0.0518	[-0.1032, -0.0004]	-2.017	0.0481	-0.512	*

Significant at $p < 0.05$: 10/12 windows

Significant at $p < 0.01$: 2/12 windows