

# Supplementary Material

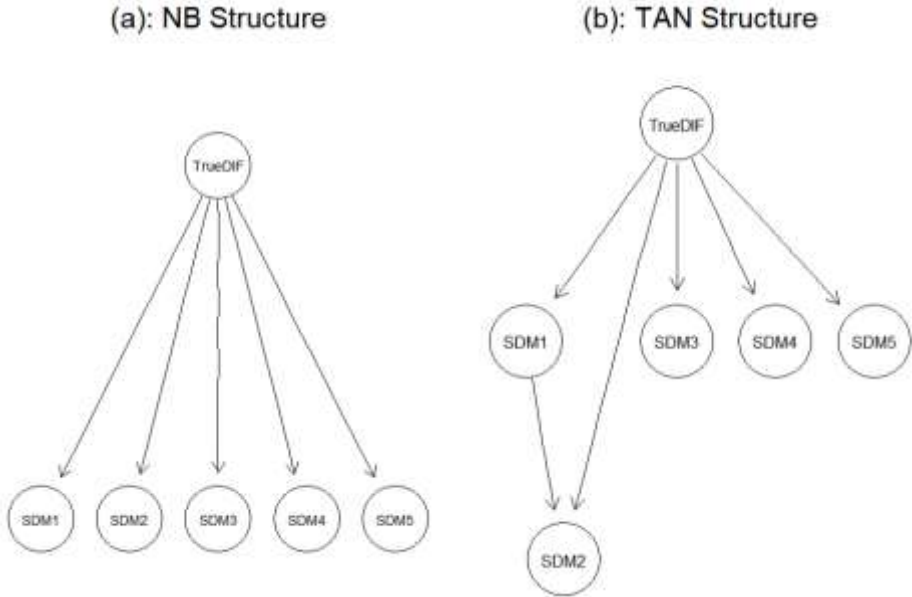


Figure S1 Examples of the NB and TAN structures

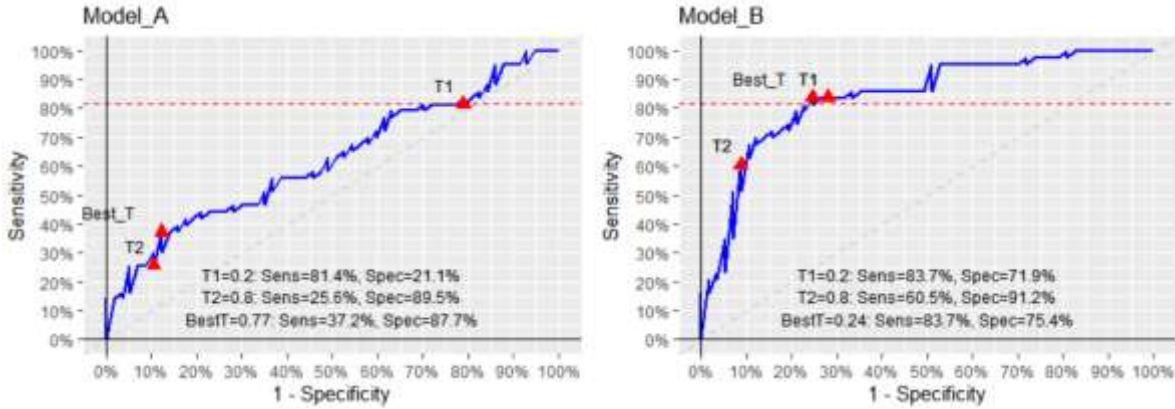
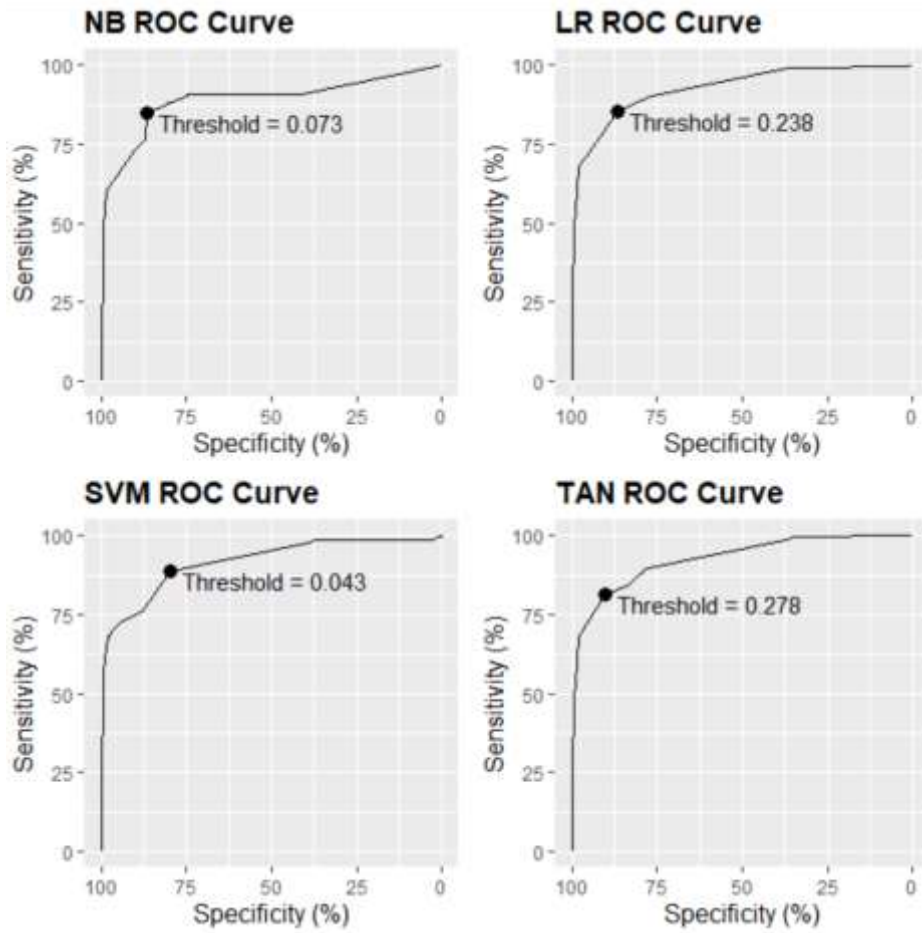


Figure S2 Example Plot of Area Under the Curve (AUC)



**Figure S3** AUC and Threshold Settings of Four MDC Models

**Table S1** Test Conditions for Simulated Data

Type of Test Conditions	Test Conditions	Level Count	Training Set (TS)&	
			Matched Test Sets (MTS)	Unmatched Test Sets (UMTS)
Observable Test Conditions (Total: 3*2*3=18)	Sample Size	3	500,1000,2000	Same as left
	Sample Size Ratio	2	R=50%	Same as left
			R=80%	Same as left
DIF-Related Test Conditions (Total: 2*2*2=8)	Test Length	3	20,40,60	Same as left
	$\Delta$ alpha	2	DIFmagnitudes_a	0,-1
			DIFmagnitudes_b	0.4,0.8
	Proportion of DIF	2	0.2,0.3	0.1,0.4
Other Unobservable Test	Impact	2	No Impact	Significant Impact:

---

Conditions

R: N (0,1);

R: N (0,1);

(Total:2)

F: N (0,1)

F: N (-1,1)

Impact

Reverse Impact:

R: N (0,1);

R: N (-0.5,1);

F: N (-0.5,1)

F: N (0,1)

aR: N(1,0.2)

aR:U(0.9,2.5)

Item parameters

-

aF=aR+Δa(DIF item)

aF=aR+Δa(DIF item)

alpha

aF=aR(noDIF item)

aF=aR(noDIF item)

bR: N(0,1)

bR:U(-1.9,1.9)

Item parameters

-

bF=bR+Δb(DIF item)

bF=bR+Δb(DIF item)

beta

bF=bR(noDIF item)

bF=bR(noDIF item)

---

**Table S2** AUC (%) for MTS Under Each Test Condition

Test Conditions	Level	Multi-Detectors Combination (MDC)					Single Detection Methods (SDM)				Simple Integration Methods (SIM)	
		LR	NB	SVM	TAN	M.H	Stand.	Logistic	Lord	Raju	Voting	AnyFlat
Samplesize	500	78.5	77.8	78.1	78.7	72.1	71.3	77.3	58.9	64.1	70.5	62.0
	1000	82.3	81.5	81.2	82.0	77.9	73.0	80.5	63.3	68.2	74.6	65.1
	2000	83.3	82.4	80.0	82.8	79.8	72.7	79.1	67.4	69.8	76.3	65.7
SamplesizeRaito	0.5	82.2	81.2	80.1	81.7	77.7	72.6	79.2	65.0	68.6	74.6	65.9
	0.8	80.5	80.0	79.5	80.6	75.4	72.0	78.8	61.4	66.1	72.9	62.7
TestLength	20	81.0	80.5	79.3	80.9	76.3	72.3	79.2	64.4	68.8	74.1	64.9
	40	81.4	80.7	80.4	81.2	76.7	72.5	78.9	62.9	67.5	73.7	64.0
	60	81.7	80.6	79.6	81.5	76.9	72.2	78.8	62.2	65.6	73.6	63.9
ProportionsofDIF	0.2	84.2	83.6	81.7	83.9	79.4	73.6	82.8	66.2	69.0	77.2	67.2
	0.3	78.5	77.6	77.8	78.4	73.8	71.0	75.2	60.1	65.7	70.3	61.3
DIFmagnitudes_a	0	73.1	73.0	68.1	72.8	72.8	63.6	72.2	66.9	64.2	70.0	69.8
	1	89.6	88.2	91.5	89.5	80.4	81.1	85.8	59.5	70.5	77.6	58.8
DIFmagnitudes_b	0.4	77.4	76.6	74.7	77.1	73.3	67.4	75.3	60.6	65.1	70.2	62.4
	0.8	85.3	84.5	84.8	85.3	79.9	77.3	82.6	65.8	69.6	77.3	66.2
Impact	R 0 F -0.5	80.7	80.0	79.4	80.7	76.1	71.8	78.2	62.4	66.2	72.9	63.6
	R 0 F 0	82.0	81.2	80.1	81.7	77.1	72.8	79.7	63.9	68.5	74.7	64.9

Note: The shaded colors in the table indicate the best results from SDM and SIM under the current test conditions.

**Table S3** Results of Repeated Measures ANOVA Analysis of AUC for MTS

Effect	SSn	SSd	DFn	DFd	F	p	$\eta^2$
Method	13538716	41422067	10	316690	10351	0.000	0.327
Method:DIFmagnitudes_a	8129019	41422067	10	316690	6215	0.000	0.196
DIFmagnitudes_a	7357765	41422067	1	316690	56253	0.000	0.178
DIFmagnitudes_b	4019250	41422067	1	316690	30729	0.000	0.097
ProportionsofDIF	2275684	41422067	1	316690	17399	0.000	0.055
Samplesize	1223780	41422067	2	316690	4678	0.000	0.030
Method:Samplesize	278871	41422067	20	316690	107	0.000	0.007
Method:DIFmagnitudes_b	292542	41422067	10	316690	224	0.000	0.007
SamplesizeRaito	237334	41422067	1	316690	1815	0.000	0.006
Method:ProportionsofDIF	162344	41422067	10	316690	124	0.000	0.004
Impact	137481	41422067	1	316690	1051	0.000	0.003
Method:SamplesizeRaito	81307	41422067	10	316690	62	0.000	0.002
Method:TestLength	82533	41422067	20	316690	32	0.000	0.002
TestLength	11547	41422067	2	316690	44	0.000	0.000
Method:Impact	14269	41422067	10	316690	11	0.000	0.000

**Table S4** AUC (%) for UMTS Under Each Test Condition

Test Conditions	Level	Multi-Detectors Combination (MDC)				Single Detection Methods (SDM)					Simple Integration Methods (SIM)	
		LR	NB	SVM	TAN	M.H	Stand.	Logistic	Lord	Raju	Voting	AnyFlat
Samplesize	500	77.1	76.8	76.5	77.3	71.8	70.0	76.5	63.0	62.7	70.7	65.0
	1000	80.6	80.4	78.8	80.7	77.0	71.3	79.8	68.4	67.3	75.3	68.2
	2000	81.5	81.1	77.5	81.4	78.6	71.7	78.5	72.0	69.7	77.6	68.2
SamplesizeRaito	0.5	80.2	79.8	77.0	80.0	76.9	71.3	78.3	69.5	68.4	75.8	68.5
	0.8	79.4	79.1	78.1	79.6	74.8	70.7	78.2	66.1	64.7	73.3	65.8
TestLength	20	79.4	79.2	76.9	79.3	75.4	70.8	78.0	69.0	67.9	74.4	68.0
	40	79.9	79.7	78.6	80.0	76.1	71.3	78.5	67.7	66.8	74.7	67.1
	60	80.0	79.5	77.1	80.1	76.0	71.0	78.3	66.6	65.0	74.4	66.3
ProportionsofDIF	0.1	85.6	85.6	81.4	85.7	81.5	73.7	85.2	74.5	71.7	81.1	74.1
	0.4	73.9	73.3	73.8	73.9	70.1	68.3	71.2	61.1	61.5	67.9	60.2
DIFmagnitudes_a	0	73.3	73.5	68.9	73.3	73.2	64.7	72.2	68.7	66.0	70.9	69.9
	1	86.2	85.3	86.2	86.3	78.4	77.2	84.2	66.9	67.2	78.1	64.4
DIFmagnitudes_b	0.4	76.9	76.9	72.8	77.0	73.6	66.1	76.0	66.4	65.3	72.4	66.5
	0.8	82.6	82.0	82.4	82.6	78.1	76.0	80.5	69.2	67.8	76.6	67.7
Impact	R -0.5 F 0	79.8	79.5	78.3	79.9	75.9	72.4	78.9	68.7	67.4	75.3	68.0
	R 0 F -1	79.7	79.4	76.9	79.7	75.8	69.6	77.6	66.8	65.7	73.8	66.2

**Note:** The shaded colors in the table indicate the best results from SDM and SIM under the current test conditions.

**Table S5** Results of Repeated Measures ANOVA Analysis of AUC for UMTS

Effect	SSn	SSd	DFn	DFd	F	p	$\eta^2$
ProportionsofDIF	10228764	50770975	1	316690	63803	0.000	0.201
Method	8064966	50770975	10	316690	5031	0.000	0.159
DIFmagnitudes_a	4878220	50770975	1	316690	30428	0.000	0.096
Method:DIFmagnitudes_a	3707935	50770975	10	316690	2313	0.000	0.073
DIFmagnitudes_b	2006143	50770975	1	316690	12514	0.000	0.040
Samplesize	1262096	50770975	2	316690	3936	0.000	0.025
Method:ProportionsofDIF	530273	50770975	10	316690	331	0.000	0.010
Method:DIFmagnitudes_b	520387	50770975	10	316690	325	0.000	0.010
Method:Samplesize	328750	50770975	20	316690	103	0.000	0.006
SamplesizeRaito	167325	50770975	1	316690	1044	0.000	0.003
Method:SamplesizeRaito	168713	50770975	10	316690	105	0.000	0.003
Impact	116194	50770975	1	316690	725	0.000	0.002
Method:TestLength	96956	50770975	20	316690	30	0.000	0.002
Method:Impact	61431	50770975	10	316690	38	0.000	0.001
TestLength	17937	50770975	2	316690	56	0.000	0.000



**Table S6** Correlations Between Various Results of SDMs and MDC (LR) (MTS)

TestConditions	Level	M.H	Stand.	Logistic	Lord	Raju	Voting	AnyFlat
Samplesize	500	0.742	0.700	0.859	0.280	-0.047	0.663	0.499
Samplesize	1000	0.849	0.598	0.817	0.358	0.005	0.645	0.497
Samplesize	2000	0.870	0.454	0.781	0.455	0.116	0.661	0.486
SamplesizeRaito	0.5	0.841	0.560	0.789	0.406	0.068	0.632	0.528
SamplesizeRaito	0.8	0.799	0.609	0.849	0.322	-0.019	0.681	0.460
TestLength	20	0.802	0.591	0.851	0.390	0.022	0.672	0.494
TestLength	40	0.829	0.576	0.807	0.352	0.015	0.647	0.491
TestLength	60	0.829	0.588	0.798	0.350	0.037	0.650	0.497
ProportionsofDIF	0.2	0.821	0.602	0.833	0.384	0.047	0.676	0.496
ProportionsofDIF	0.3	0.819	0.568	0.805	0.344	0.001	0.636	0.492
DIFmagnitudes_a	0	0.833	0.507	0.856	0.585	0.409	0.703	0.712
DIFmagnitudes_a	1	0.808	0.661	0.781	0.144	-0.359	0.609	0.273
DIFmagnitudes_b	0.4	0.804	0.543	0.808	0.349	0.006	0.629	0.487
DIFmagnitudes_b	0.8	0.836	0.626	0.829	0.380	0.042	0.684	0.501
Impact	R:0F:-0.5	0.820	0.578	0.817	0.346	0.014	0.650	0.484
Impact	R:0F:0	0.820	0.592	0.821	0.383	0.034	0.663	0.504

**Table S7** Correlations Between Various Results of SDMs and MDC (LR) (UMTS)

TestConditions	Level	M.H	Stand.	Logistic	Lord	Raju	Voting	AnyFlat
Samplesize	500	0.724	0.638	0.851	0.448	0.202	0.653	0.609
Samplesize	1000	0.821	0.510	0.811	0.483	0.230	0.631	0.595
Samplesize	2000	0.845	0.428	0.738	0.562	0.317	0.646	0.552
SamplesizeRaito	0.5	0.813	0.520	0.748	0.520	0.280	0.613	0.584
SamplesizeRaito	0.8	0.781	0.531	0.852	0.476	0.219	0.673	0.587
TestLength	20	0.756	0.518	0.824	0.528	0.256	0.628	0.612
TestLength	40	0.811	0.529	0.782	0.493	0.245	0.636	0.584
TestLength	60	0.823	0.530	0.794	0.473	0.248	0.666	0.560
ProportionsofDIF	0.1	0.790	0.566	0.811	0.525	0.315	0.666	0.600
ProportionsofDIF	0.4	0.803	0.485	0.789	0.471	0.184	0.620	0.570
DIFmagnitudes_a	0	0.800	0.463	0.818	0.602	0.432	0.666	0.704
DIFmagnitudes_a	1	0.793	0.588	0.782	0.393	0.067	0.620	0.466
DIFmagnitudes_b	0.4	0.782	0.497	0.801	0.503	0.262	0.638	0.603
DIFmagnitudes_b	0.8	0.811	0.554	0.799	0.493	0.237	0.648	0.567
Impact	R:-0.5F:0	0.824	0.562	0.830	0.552	0.321	0.685	0.622
Impact	R:0F:-1	0.770	0.489	0.770	0.444	0.178	0.601	0.549

**Table S8** Performance of Models on Real Data

Performance	MDC				SDM				
	NB	LR	SVM	TAN	M.H	Std	Logistic	Lord	Raju
AUC(%)	85.5	85.1	83.7	84.2	80.3	74.9	84.1	65.0	55.2
Specificity (%)	84.0	84.6	88.3	78.8	77.8	50.5	89.2	81.4	58.0
Sensitivity (%)	87.1	85.6	79.1	89.6	82.8	99.4	78.9	48.6	52.3