

Results for Pre-Proc Learning Fair Representations:

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Disparate Impact (Before)	1.0391294	1.03474102	0.3620832	0.8648391
Disparate Impact (After)	1.0093853	0.99841252	0.1198747	1.0021945

Logistic Regression				
Accuracy (Before)	0.8091111	0.81546078	0.8195311	0.7500000
Accuracy (After)	0.7765555	0.78642961	0.7582147	0.5100000
Balanced Accuracy (Before)	0.6432656	0.63359499	0.6914758	0.6155555
Balanced Accuracy (After)	0.5055595	0.50606092	0.5004230	0.6111111

Random Forest				
Accuracy (Before)	0.8013333	0.81174749	0.8344764	0.7633333
Accuracy (After)	0.7758888	0.78642961	0.7581123	0.5033333
Balanced Accuracy (Before)	0.5963679	0.60121466	0.6781526	0.5711111
Balanced Accuracy (After)	0.5044285	0.50454117	0.5002115	0.5755555

Statistical Parity (Before)	0.0298339	0.02647544	0.1929391	0.0956693
Statistical Parity (After)	0.0092356	0.00154118	0.0085174	0.0011153

Consistency (Before)	0.7854190	0.78197338	0.8326312	0.6968571
Consistency (After)	0.9996666	0.99951775	0.9997287	0.9720000

Table showing Logistic Regression and Random Forest Classification (RFC) model performance when attempting to predict the Sensitive class from the transformed data

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Logistic Regression - Sensitive attribute model performance				
Accuracy (Before)	0.6068888	0.70653763	0.7607738	0.7133333
Accuracy (After)	0.5178888	0.48115224	0.3532603	0.6800000
Balanced Accuracy (Before)	0.5373862	0.69811529	0.6952920	0.5851715
Balanced Accuracy (After)	0.4857890	0.50641664	0.4790547	0.5000000

Random Forest - Sensitive attribute model performance				
Accuracy (Before)	0.6177777	0.73320580	0.8003889	0.7100000
Accuracy (After)	0.6003333	0.50849555	0.3227556	0.6666666
Balanced Accuracy (Before)	0.5205672	0.73264158	0.7569811	0.5634191
Balanced Accuracy (After)	0.4983500	0.52825630	0.4362819	0.4901960

		Taiwan Gender	Taiwan Marriage	Adult	German
Training	Before Label=0	4611	4636	29750	344
	After Label = 0	218	622	38805	225
	Before Label=1	16389	16100	9322	356
	After Label = 1	20782	20114	268	475
		(21000)	(20736)	(39073)	(700)

Results for Pre-Proc Disparate Impact Removal:

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Disparate Impact (Before)	1.0437451	1.03520733	0.3620832	0.8722222
Disparate Impact (After)	1.1722253	0.96055666	0.7838038	0.8270042
Repair Level	0.5	0.1	1.0	0.5

Logistic Regression				
Accuracy (Before)	0.8083333	0.81620253	0.8195311	0.7200000
Accuracy (After)	0.5390000	0.55021097	0.7376394	0.6600000
Balanced Accuracy (Before)	0.6481056	0.64200441	0.6914758	0.5742523
Balanced Accuracy (After)	0.6196062	0.62180424	0.7330637	0.6842510

Random Forest				
Accuracy (Before)	0.7946666	0.81299578	0.8344764	0.7600000
Accuracy (After)	0.7996666	0.81400843	0.8381615	0.7400000
Balanced Accuracy (Before)	0.5845750	0.60384957	0.6781526	0.5592350
Balanced Accuracy (After)	0.6037297	0.60191430	0.7092362	0.5395969

Statistical Parity (Before)	0.0332706	0.02685611	0.1929391	0.0919999
Statistical Parity (After)	0.0666004	0.01790784	0.0864186	0.0976190

Consistency (Before)	0.7844583	0.78298591	0.8326312	0.7055000
Consistency (After)	0.9187666	0.91858227	0.9120073	0.7599999

Table showing Logistic Regression and Random Forest Classification (RFC) model performance when attempting to predict the Sensitive class from the transformed data

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Logistic Regression - Sensitive attribute model performance				
Accuracy (Before)	0.6075000	0.71105485	0.7607738	0.7350000
Accuracy (After)	0.6073333	0.71172995	0.6649605	0.7300000
Balanced Accuracy (Before)	0.5391934	0.70077261	0.8003889	0.7100000
Balanced Accuracy (After)	0.5390558	0.70175749	0.7480806	0.6071428

Random Forest - Sensitive attribute model performance				
Accuracy (Before)	0.6060000	0.73856540	0.6952920	0.6202380
Accuracy (After)	0.6056666	0.73620253	0.5509809	0.7250000
Balanced Accuracy (Before)	0.5121004	0.73806391	0.7569811	0.5595238
Balanced Accuracy (After)	0.5004226	0.73762533	0.6635963	0.5750000

		Taiwan Gender	Taiwan Marriage	Adult	German
Training	Before Label=0	5265	5278	29750	247
	After Label = 0	5265	5278	29750	553
	Before Label=1	18735	18420	9323	247
	After Label = 1	18735	18420	9323	553

Results for Pre-Proc Reweighing:

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Disparate Impact (Before)	1.0467924	1.03252337	0.3685749	0.9130025
Disparate Impact (After)	1.0000000	1.00000000	1.0000000	0.9999999

Logistic Regression				
Accuracy (Before)	0.8110000	0.81687763	0.8229092	0.6950000
Accuracy (After)	0.8100000	0.81940928	0.8147200	0.7000000
Balanced Accuracy (Before)	0.5941323	0.64975414	0.6861426	0.5888372
Balanced Accuracy (After)	0.5948680	0.65477351	0.6616281	0.5960836

Random Forest				
Accuracy (Before)	0.8118333	0.80084388	0.8285392	0.6900000
Accuracy (After)	0.8123333	0.80101265	0.8234210	0.6750000
Balanced Accuracy (Before)	0.5982413	0.58712446	0.6664366	0.5575838
Balanced Accuracy (After)	0.6004866	0.58749374	0.6469660	0.5427038

Statistical Parity (Before)	0.0354055	0.02495861	0.1906723	0.0636488
Statistical Parity (After)	2.220e-16	0.00000000	8.326e-17	2.220e-16

Consistency (Before)	0.7836833	0.7831125	0.8323445	0.7037500
Consistency (After)	0.7836833	0.7831125	0.8323445	0.7037500

Table showing Logistic Regression and Random Forest Classification (RFC) model performance when attempting to predict the Sensitive class from the transformed data

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Logistic Regression - Sensitive attribute model performance				
Accuracy (Before)	0.6063333	0.73552742	0.7693725	0.7700000
Accuracy (After)	0.6095000	0.73839662	0.7685535	0.7700000
Balanced Accuracy (Before)	0.5379510	0.73714819	0.7019187	0.6149162
Balanced Accuracy (After)	0.5379510	0.73714819	0.7019187	0.6149162

Random Forest - Sensitive attribute model performance				
Accuracy (Before)	0.6128333	0.73215189	0.8007984	0.7650000
Accuracy (After)	0.6121666	0.73248945	0.8023339	0.7600000
Balanced Accuracy (Before)	0.5251818	0.73518748	0.7566385	0.5706494
Balanced Accuracy (After)	0.5251818	0.73518748	0.7566385	0.5706494

		Taiwan Gender	Taiwan Marriage	Adult	German
Whole ds	Before Label=0	6636	6547	37115	300
	After Label = 0	6636	6547	37115	300
	Before Label=1	23364	23076	11687	700
	After Label = 1	23364	23076	11687	700

ONLY WEIGHTS ARE APPLIED.

Results for In-Proc : Adversarial Debiasing (using training dataset)

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Disparate Impact (Before)	1.0409955	1.03057494	0.3616879	0.9028143
Disparate Impact (After)	0.9662387	1.01510716	0.7974075	0.4485465
Repair Level				

Logistic Regression				
Accuracy (Before)	0.8216666	0.81535864	0.8217831	0.780000
Accuracy (After)	0.9496666	0.97383966	0.9034701	0.930000
Balanced Accuracy (Before)	0.6470951	0.60022919	0.6818849	0.6331096
Balanced Accuracy (After)	0.8460778	0.91581243	0.7480738	0.8762183

Random Forest				
Accuracy (Before)	0.8121666	0.81147679	0.8353976	0.750000
Accuracy (After)	0.9203333	0.94092827	0.9175964	0.920000
Balanced Accuracy (Before)	0.6002443	0.58665843	0.6725080	0.5484932
Balanced Accuracy (After)	0.7485771	0.78915013	0.7403673	0.7995451

Statistical Parity (Before)	0.0310865	0.02335915	0.1942230	0.0690158
Statistical Parity (After)	0.0308645	0.01319203	0.0343408	0.5254791

Consistency (Before)	0.7812333	0.78331505	0.8316945	0.694750
Consistency (After)	0.9738083	0.97370242	0.9591533	0.890250

Table showing Logistic Regression and Random Forest Classification (RFC) model performance when attempting to predict the Sensitive class from the transformed data

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Logistic Regression - Sensitive attribute model performance				
Accuracy (Before)	0.6143333	0.72590717	0.7548367	0.765000
Accuracy (After)	0.6285000	0.72708860	0.7636400	0.850000
Balanced Accuracy (Before)	0.5303988	0.72781386	0.6893517	0.6431743
Balanced Accuracy (After)	0.5556714	0.72913436	0.6898084	0.7669471

Random Forest - Sensitive attribute model performance				
Accuracy (Before)	0.6151666	0.72118143	0.7975227	0.700000
Accuracy (After)	0.6155000	0.72202531	0.7946565	0.835000
Balanced Accuracy (Before)	0.5112115	0.72774844	0.7554315	0.5250116
Balanced Accuracy (After)	0.5089027	0.72852605	0.7525725	0.7516362

Table showing label value change before and after model transformation

		Taiwan Gender	Taiwan Marriage	Adult	German
Training	Before Label=0	5352	5295	29713	249
	After Label = 0	2505	2836	32897	168
	Before Label=1	18648	18403	9360	551
	After Label = 1	21495	20862	6176	632
Test	Before Label=0	1284	1252	7442	51
	After Label = 0	5064	826	8247	38
	Before Label=1	4716	4673	2327	149
	After Label = 1	936	5099	1522	162

Results for Post-Proc Reject Option Classification:

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Disparate Impact (Before)	1.0691643	1.04752642	0.3605318	1.0714285
Disparate Impact – Log (After)	1.0504995	1.10111443	0.6281866	1.0129870
Disparate Impact – Rfc (After)	1.0232405	1.02622463	0.5079544	1.3775510

Logistic Regression				
Accuracy (Before)	0.8070000	0.80787408	0.8198756	0.7450000
Accuracy (After)	0.7760833	0.77626803	0.8234842	0.7337500
Balanced Accuracy (Before)	0.6007241	0.60223492	0.6783533	0.6068437
Balanced Accuracy (After)	0.6889000	0.69090000	0.6824996	0.6314000

Random Forest				
Accuracy (Before)	0.8096666	0.80998396	0.8344892	0.7625000
Accuracy (After)	0.8096666	0.80998396	0.8344892	0.7612500
Balanced Accuracy (Before)	0.6180156	0.61884401	0.6709555	0.6553000
Balanced Accuracy (After)	0.7264000	0.71930000	0.6709555	0.6948000

Statistical Parity (Before)	0.0519715	0.03613819	0.2003601	0.0476190
Statistical Parity Log (After)	0.0399190	0.07523354	0.0350745	0.0060000
Statistical Parity Rfc (After)	0.0167206	0.01838759	0.0370729	0.1468253

Consistency (Before)	0.7388666	0.73857577	0.8245240	0.6640000
Consistency Log (After)	0.7840000	0.76361795	0.9578710	0.6140000
Consistency Rfc (After)	0.7476000	0.74114074	0.9882906	0.6300000

Table showing Logistic Regression and Random Forest Classification (RFC) model performance when attempting to predict the Sensitive class from the transformed data

	Taiwan-Gender	Taiwan-Marriage	Adult	German
Logistic Regression - Sensitive attribute model performance				
Accuracy (Before)	0.5065000	0.64168776	0.7626164	0.7750000
Accuracy (After)	0.5065000	0.64168776	0.7626164	0.7750000
Balanced Accuracy (Before)	0.5008655	0.64191053	0.6957168	0.6416927
Balanced Accuracy (After)	0.5008655	0.64191053	0.6957168	0.6416927

Random Forest - Sensitive attribute model performance				
Accuracy (Before)	0.5156666	0.73400843	0.8007984	0.7600000
Accuracy (After)	0.5156666	0.73400843	0.8007984	0.7600000
Balanced Accuracy (Before)	0.5102468	0.73508547	0.7531003	0.5805642
Balanced Accuracy (After)	0.5102468	0.73508547	0.7531003	0.5805642

Test		Taiwan Gender	Taiwan Marriage	Adult	German
LogReg	Before Label=0	663	656	3679	32
	After Label = 0	565	646	4481	46
	Before Label=1	2337	2307	1206	68
	After Label = 1	2435	2317	404	54
RFC	Before Label=0	663	656	3679	32
	After Label = 0	815	858	4577	43
	Before Label=1	2337	2307	1206	68
	After Label = 1	2185	2105	308	57