

Gender Fairness Results

April 2025

1 Results

Gender	Users	Tracks	Artists	LEs
ALL	39156	10000	7005	4728342
F	10643	8465	5992	1190536
M	28513	9993	7001	3537806

Table 1: Final Data Statistics

Model	Scenario	NDCG@10	Recall@10	Diversity@10	Coverage@10
POP	Standard	0.053	0.051	0.927	0.002
	Resampled	0.034	0.034	0.925	0.002
	SMOTE	0.052	0.050	0.927	0.002
ItemKNN	Standard	0.162	0.127	0.935	0.830
	Resampled	0.128	0.097	0.936	0.831
	SMOTE	0.144	0.111	0.935	0.831
SLIM _{100k}	Standard	0.068	0.085	0.983	0.561
	Resampled	0.055	0.072	0.982	0.581
	SMOTE	0.058	0.068	0.982	0.538
SLIM _{100k-Distribution Aware}	Standard	0.068	0.085	0.983	0.561

Table 2: Overall results of accuracy (NDCG and recall) and beyond-accuracy (diversity and coverage) metrics on the RS algorithms

Model	Scenario	ALL	M/F	RecGap	Score Dist. (M/F)	CompFct
POP	Standard	0.052	0.047/0.065	0.018 (f)	69.2/30.8	0.011
	Resampled	0.034	0.043/0.009	0.034 (m)	82.6/17.4	0.276
	SMOTE	0.053	0.047/0.061	0.013 (f)	43.8/56.2	0.023
ItemKNN	Standard	0.162	0.166/0.152	0.014 (m)	77.3/22.7	0.001
	Resampled	0.128	0.166/0.030	0.136 (m)	84.7/15.3	0.329
	SMOTE	0.144	0.136/0.155	0.018 (f)	46.8/53.2	0.029
SLIM _{100k}	Standard	0.068	0.071/0.059	0.012 (m)	78.9/21.1	0.003
	Resampled	0.055	0.055/0.052	0.003 (m)	51.4/48.6	0.031
	SMOTE	0.058	0.047/0.075	0.027 (f)	38.9/61.0	0.026
SLIM _{100k-Distribution Aware}	Standard	0.068	0.071/0.059	0.012 (m)	78.9/21.1	0.003

Table 3: NDCG@10 results on the users of the male/female (M/F) groups for the Standard and Resampled scenarios.

Model	Scenario	ALL	M/F	RecGap	Score Dist. (M/F)	CompFct
POP	Standard	0.051	0.044/0.069	0.024 (f)	66.8/33.2	0.019
	Resampled	0.034	0.046/0.004	0.042 (m)	92.1/7.9	0.617
	SMOTE	0.050	0.046/0.056	0.010 (f)	45.2/54.8	0.026
ItemKNN	Standard	0.127	0.127/0.125	0.002 (m)	76.0/24.0	0.000
	Resampled	0.097	0.130/0.012	0.118 (m)	91.8/8.2	0.601
	SMOTE	0.111	0.108/0.118	0.010 (f)	47.7/52.3	0.030
SLIM _{100k}	Standard	0.085	0.085/0.082	0.003 (m)	76.4/23.6	0.000
	Resampled	0.072	0.073/0.069	0.004 (m)	51.3/48.7	0.032
	SMOTE	0.068	0.066/0.073	0.007 (f)	47.7/52.3	0.001
SLIM _{100k-Distribution Aware}	Standard	0.085	0.085/0.082	0.003 (m)	76.4/23.6	0.000

Table 4: Recall@10 results on the users of the male/female (M/F) groups for the Standard and Resampled scenarios.

Model	Scenario	ALL	M/F	RecGap	Score Dist. (M/F)	CompFct
POP	Standard	0.927	0.927/0.927	0.000 (f)	75.7/24.3	0.000
	Resampled	0.925	0.925/0.925	0.000 (f)	50.0/50.0	0.000
	SMOTE	0.927	0.927/0.927	0.000 (f)	50.0/50.0	0.000
ItemKNN	Standard	0.936	0.935/0.937	0.002 (f)	75.7/24.3	0.000
	Resampled	0.936	0.936/0.938	0.002 (f)	50.0/50.0	0.000
	SMOTE	0.935	0.935/0.935	0.000 (m)	50.0/50.0	0.000
SLIM _{100k}	Standard	0.983	0.983/0.984	0.001 (f)	75.7/24.3	0.000
	Resampled	0.982	0.982/0.982	0.000 (m)	50.0/50.0	0.000
	SMOTE	0.982	0.981/0.982	0.001 (f)	50.2/49.8	0.000
SLIM _{100k-Distribution Aware}	Standard	0.983	0.983/0.984	0.001 (f)	75.7/24.3	0.000

Table 5: Diversity@10 results on the users of the male/female (M/F) groups for the Standard and Resampled scenarios.

Model	Scenario	ALL	M/F	RecGap	Score Dist. (M/F)	CompFct
POP	Standard	0.002	0.002/0.002	0.000 (f)	74.6/25.4	0.000
	Resampled	0.002	0.002/0.002	0.000 (m)	52.9/47.1	0.002
	SMOTE	0.002	0.002/0.002	0.000 (f)	50.0/50.0	0.031
ItemKNN	Standard	0.830	0.782/0.534	0.248 (m)	82.1/17.9	0.012
	Resampled	0.831	0.782/0.520	0.262 (m)	60.1/39.9	0.021
	SMOTE	0.831	0.760/0.609	0.152 (m)	55.5/44.5	0.025
SLIM _{100k}	Standard	0.561	0.500/0.295	0.205 (m)	84.1/15.9	0.023
	Resampled	0.581	0.512/0.316	0.196 (m)	61.8/38.2	0.004
	SMOTE	0.538	0.465/0.307	0.158 (m)	60.4/39.6	0.021
SLIM _{100k-Distribution Aware}	Standard	0.561	0.500/0.295	0.205 (m)	84.1/15.9	0.023

Table 6: Coverage@10 results on the users of the male/female (M/F) groups for the Standard and Resampled scenarios.