

Aggregated Consistency Metrics (Mean & Standard Deviation)

XAI Method	Jaccard Similarity Mean (std)	Spearman's ρ Mean (std)	Kendall's τ_b Mean (std)
LIME	0.835 (0.193)	0.851 (0.206)	0.776 (0.235)
SHAP	1.000 (0.000)	1.000 (0.000)	1.000 (0.000)

LIME Top-5 Feature Contributions Across Different Seeds From Explanations

id	seed	lime_top5
18694	3279	[('retarded', -0.530), ('retards', -0.296), ('world', 0.050), ('predisposes', 0.042), ('their', 0.041)]
18694	14593	[('retarded', -0.508), ('retards', -0.319), ('their', 0.054), ('killing', -0.041), ('and', 0.038)]
18694	83811	[('retarded', -0.532), ('retards', -0.299), ('their', 0.052), ('predisposes', 0.039), ('world', 0.038)]
18694	97197	[('retarded', -0.515), ('retards', -0.308), ('world', 0.057), ('their', 0.053), ('predisposes', 0.045)]

SHAP Top-5 Feature Contributions Across Different Seeds From Explanations

id	seed	shap_top5
18694	3279	[('retarded', -0.245), ('ret', -0.244), ('ards', -0.243), ('animals', -0.116), ('killing', -0.097)]
18694	14593	[('retarded', -0.245), ('ret', -0.244), ('ards', -0.243), ('animals', -0.116), ('killing', -0.097)]
18694	83811	[('retarded', -0.245), ('ret', -0.244), ('ards', -0.243), ('animals', -0.116), ('killing', -0.097)]
18694	97197	[('retarded', -0.245), ('ret', -0.244), ('ards', -0.243), ('animals', -0.116), ('killing', -0.097)]