Testing Multiple Iterations - Summary Stats

Zero shot learning

One shot learning

Table 1: Performance of model in zero-shot learning across 100 classifications of each comment at a temperature of 0.3

	Metric	Mean (%)	Standard Error (%)
Racism	Accuracy	58	6.5
	Precision	58	6.7
	Recall	59	9.2
	F1	58	6.7
Sexism	Accuracy	55	5.2
	Precision	53	3.7
	Recall	79	6.9
	F1	63	4.3
Overall	Accuracy	56	4.3
	Precision	55	3.5
	Recall	69	5.9
	F1	70	5.7

Table 2: Performance of model in one-shot learning across 100 classifications of each comment at a temperature of 0.3

	Metric	Mean (%)	Standard Error (%)
Racism	Accuracy	62	1.1
	Precision	58	0.8
	Recall	80	1.2
	F1	67	0.9
Sexism	Accuracy	60	1.4
	Precision	57	1.1
	Recall	75	1.7
	F1	65	1.2
Overall	Accuracy	61	0.9
	Precision	58	0.7
	Recall	77	1.0
	F1	64	1.3