



## Parameter Sensitivity Analysis

This section quantifies how CIF performance varies with key configuration parameters, enabling practitioners to calibrate defenses for their specific deployment contexts.

**Reproducibility:** All sensitivity data generated by `scripts/run_sensitivity_analysis.py` → `output/data/sensitivity_results.json`.

### Firewall Threshold Sensitivity

Table 1: Firewall threshold sensitivity analysis.

$\tau_{firewall}$	TPR	95% CI	FPR	95% CI	F1
0.3	0.98	[0.96, 0.99]	0.18	[0.15, 0.22]	0.90
0.4	0.97	[0.95, 0.98]	0.12	[0.09, 0.15]	0.93
0.5	0.94	[0.92, 0.96]	0.06	[0.04, 0.08]	0.94
0.6	0.91	[0.88, 0.93]	0.04	[0.02, 0.06]	0.93
0.7	0.87	[0.84, 0.90]	0.02	[0.01, 0.04]	0.92
0.8	0.82	[0.78, 0.85]	0.01	[0.00, 0.02]	0.90
0.9	0.72	[0.67, 0.76]	0.01	[0.00, 0.02]	0.84

**Optimal threshold:**  $\tau^* = 0.5$  maximizes F1 score.