

# GoEmotions: Raw vs Simplified — Model Comparison

2025-10-13 18:44 UTC

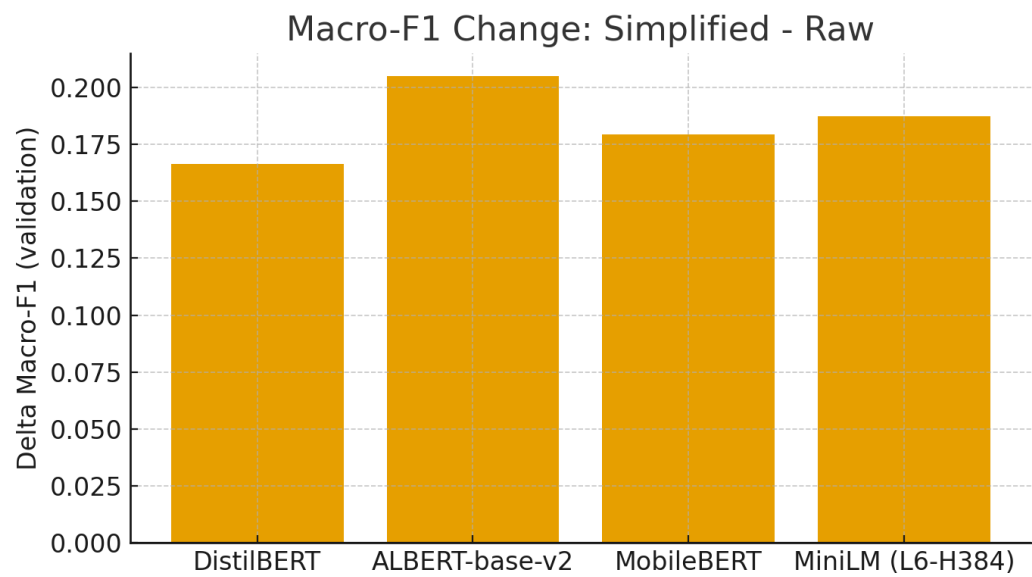
## Evaluation Metrics (Macro-F1, Validation/Test)

Model	Raw Val Macro	Simp Val Macro	Raw Test Macro	Simp Test Macro
DistilBERT	0.291	0.457	0.291	0.459
ALBERT-base-v2	0.256	0.461	0.256	0.459
MobileBERT	0.282	0.461	0.282	0.461
MiniLM (L6-H384)	0.162	0.349	0.162	0.344

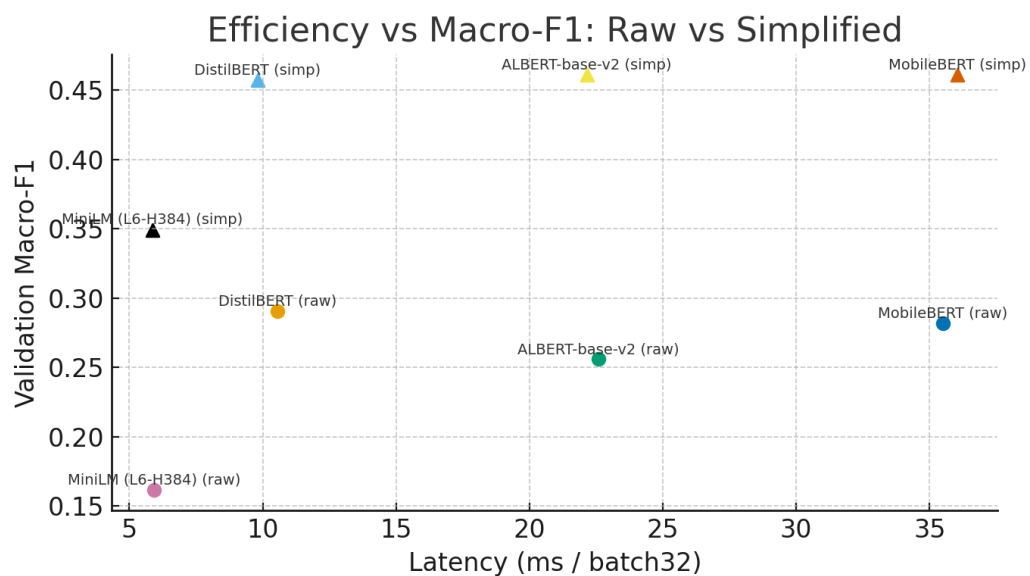
## Efficiency Snapshot

Model	Raw Params (M)	Simp Params (M)	Raw Latency	Simp Latency
DistilBERT	66.98	66.98	10.55	9.82
ALBERT-base-v2	11.71	11.71	22.59	22.15
MobileBERT	24.60	24.60	35.51	36.05
MiniLM (L6-H384)	22.72	22.72	5.92	5.86

## Macro-F1 Change (Simplified - Raw)



## Latency vs Macro-F1 (Raw vs Simplified)



## Notes

### Notes:

- Raw: 28-class multi-label (sigmoid + optimized threshold).
- Simplified: 28-class single-label (argmax).
- Macro-F1 values shown for validation and test; bold = higher within a model, italics = lower.
- Efficiency from snapshots; scatter uses circles (raw) and triangles (simplified).