

GoEmotions Baseline Report

2025-09-29 05:04 UTC

Run Summary

Model: distilbert-base-uncased
Dataset: go_emotions (raw)
Training: epochs=3, batch_size=32, max_length=128, threshold=0.5

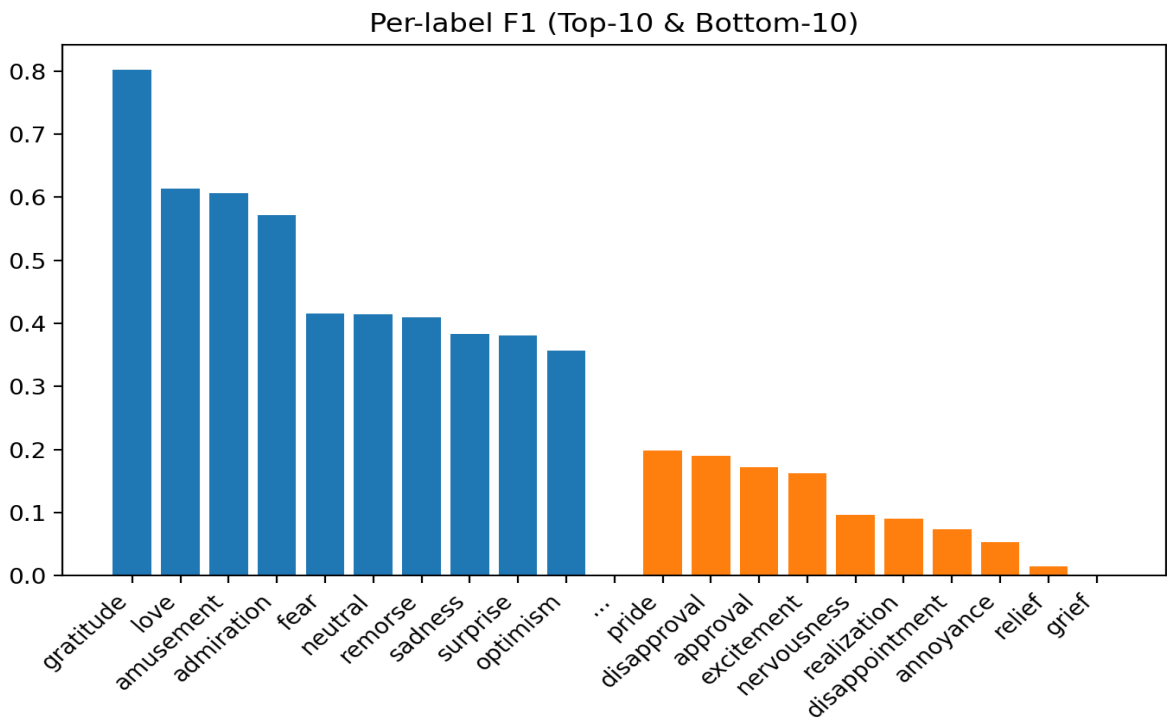
Evaluation Metrics

	F1 (micro)	F1 (macro)	F1 (weighted)	Loss
Validation	0.3759	0.2906	0.3437	0.1101
Test	0.3759	0.2906	0.3437	0.1101

Efficiency Snapshot

Trainable parameters: 66,975,004
Avg latency (ms) per batch of 32: 9.29

Per-label F1 — Top & Bottom 10



Notes

- Notes:**
- Fixed decision threshold by default. Run the threshold sweep cell to optimize Macro-F1.
 - Latency is a simple forward pass benchmark; serving latency varies by hardware and batch size.