Multilingual Polarization Detection for SemEval-25 Task 11

Ankit Dash (24BCS016) Priyanshu Mittal (24BDS058)

Piyush Prashant (24BDS055)

November 16, 2025

Abstract

Background: Detecting polarized content in multilingual social media is critical for understanding online discourse and mitigating societal divisions. This work addresses SemEval 2025 Task 11, Subtask 1: Binary Polarization Detection across 13 typologically diverse languages including English, Arabic, German, Spanish, Italian, Urdu, Chinese, Hausa, Amharic, Turkish, Hindi, Nepali, and Persian.

Methods: This research presents a production-ready multilingual polarization detection system built on XLM-RoBERTa-base with advanced training optimizations. The training pipeline incorporates stratified data splitting (85/15) maintaining both class and language balance across 40,395 training samples. Key architectural innovations include balanced class weighting computed per-dataset for optimal F1 optimization, mixed precision training (AMP) for 2x memory efficiency, gradient accumulation (2x steps) enabling larger effective batch sizes, early stopping with patience=3 to prevent overfitting, and linear warmup (6%) with decay scheduling for stable convergence. The model employs a 256-token context window, AdamW optimizer with $\beta_2 = 0.98$ for multilingual stability, and per-language F1 evaluation for fine-grained performance tracking.

Evolution: This work represents the culmination of a 24-day iterative development process (October 24 - November 16, 2025) spanning six distinct architectural phases. Phase 1 established the foundation with BitNet 1.58-bit quantization on BERT, achieving F1 Macro 0.977 on English validation. Phase 2 scaled to 9 languages with mDeBERTa-v3, achieving multilingual F1 Macro 0.764. Phase 3 introduced data augmentation with Easy Data Augmentation (EDA) and LoRA adapters for language-specific tuning. Phase 4 explored efficiency with RWKV $O(N)$ architecture, achieving 2x training speedup and 30% memory reduction while maintaining competitive accuracy. Phase 5 investigated experimental Mamba state-space models as transformer alternatives. Phase 6 delivered the production-ready XLM-RoBERTa pipeline with comprehensive error handling, model versioning, and deployment infrastructure.

Results: The final model achieved validation F1 Macro 0.7876 and accuracy 0.7876 on the held-out validation set, demonstrating robust cross-lingual transfer. Training converged in 3 epochs with early stopping triggered at epoch 6, preventing overfitting. The system processed 34,335 training samples across 13 languages with automatic checkpoint management and per-language performance monitoring. Class imbalance was effectively handled through computed class weights (0: 1.0419, 1: 0.9613), ensuring balanced learning across polarized and non-polarized classes.

Significance: This work demonstrates that production-grade multilingual NLP systems can be developed through systematic architectural evolution and careful engineering of training infrastructure. The repository provides six distinct model variants—from efficient BitNet quantization to state-of-the-art transformer architectures—enabling researchers to select appropriate trade-offs between accuracy, speed, and resource constraints. The comprehensive training pipeline with stratified sampling, mixed precision, gradient accumulation, and early stopping estab-

lishes best practices for multilingual classification tasks. Future work includes hierarchical reasoning models with multi-task learning frameworks, LLM-based approaches with parameter-efficient fine-tuning, and explainability mechanisms for human-interpretable polarization detection.