

mBART Multilingual Classification Report

Generated on: 22 Dec 2025, 12:44

Model Information

Model: mBART (encoder–decoder used for sequence classification)

Task: Multilingual Hate-Speech Classification

Data Split: 70% Train / 10% Validation / 30% Test

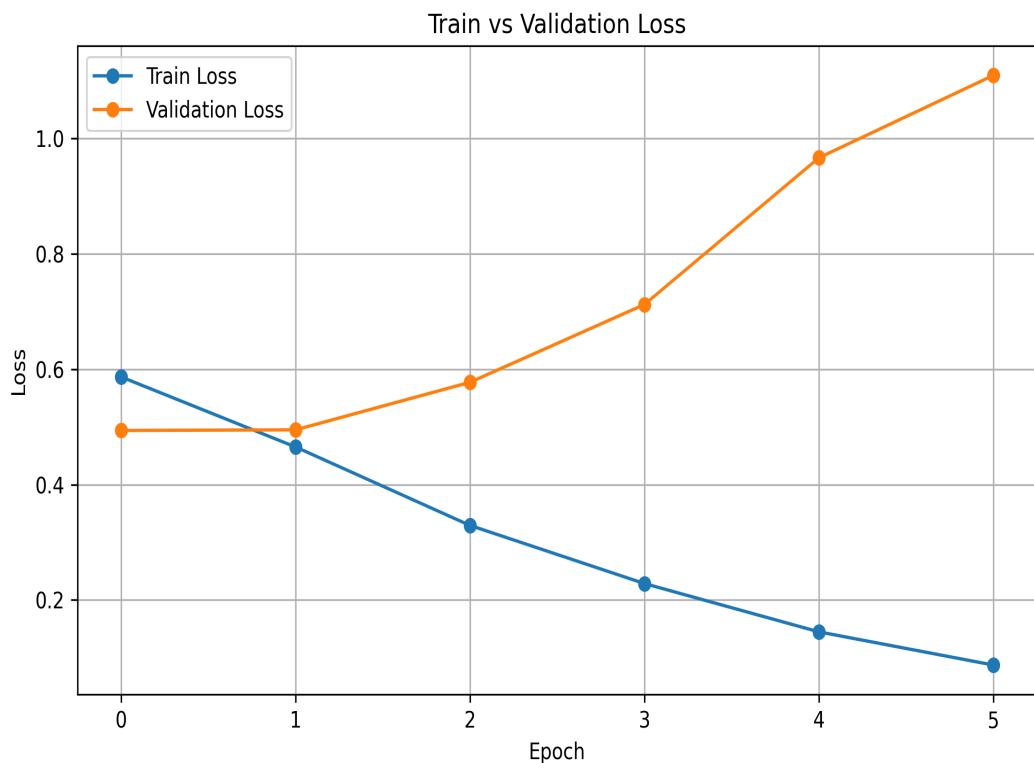
Training Hyperparameters

Parameter	Value
Learning Rate	2e-05
Batch Size	32
Epochs	8
Optimizer	AdamW
Scheduler	Linear + Warmup
Warmup Ratio	0.1
Max Sequence Length	128
Early Stopping	Patience = 2 (Val F1)

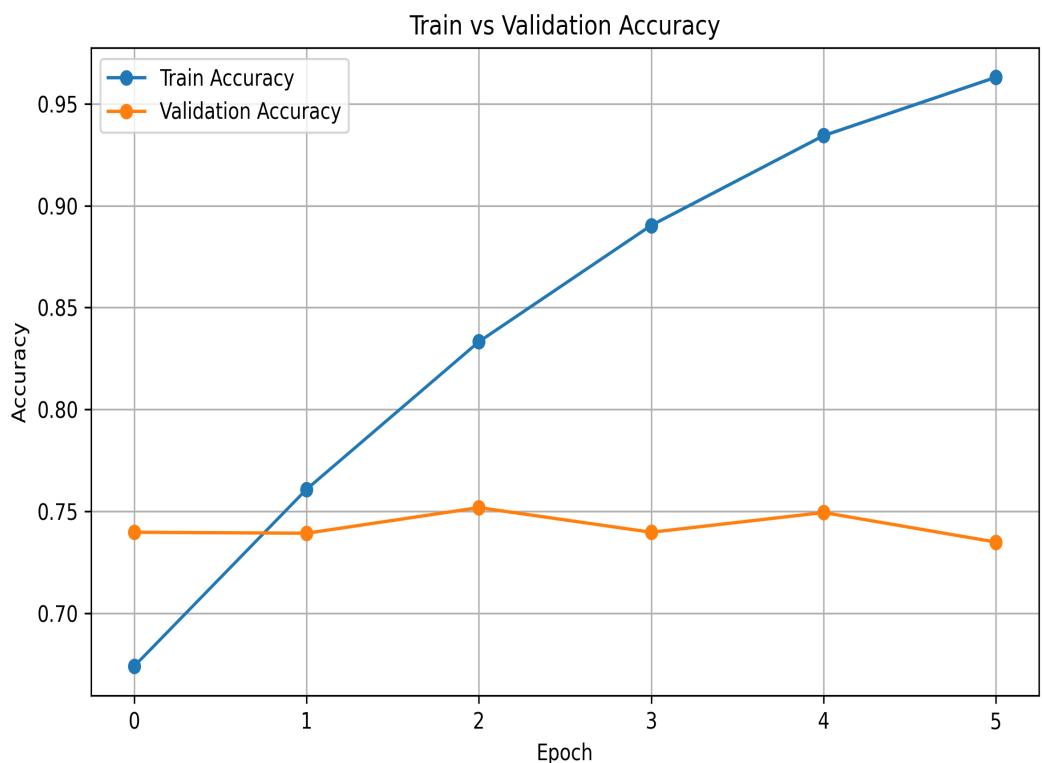
Test Set Evaluation Metrics

Metric	Value
Accuracy	0.7299
Balanced Accuracy	0.7307
Precision	0.6967
Recall	0.7411
Specificity	0.7203
F1-score	0.7182
AUC-ROC	0.8217

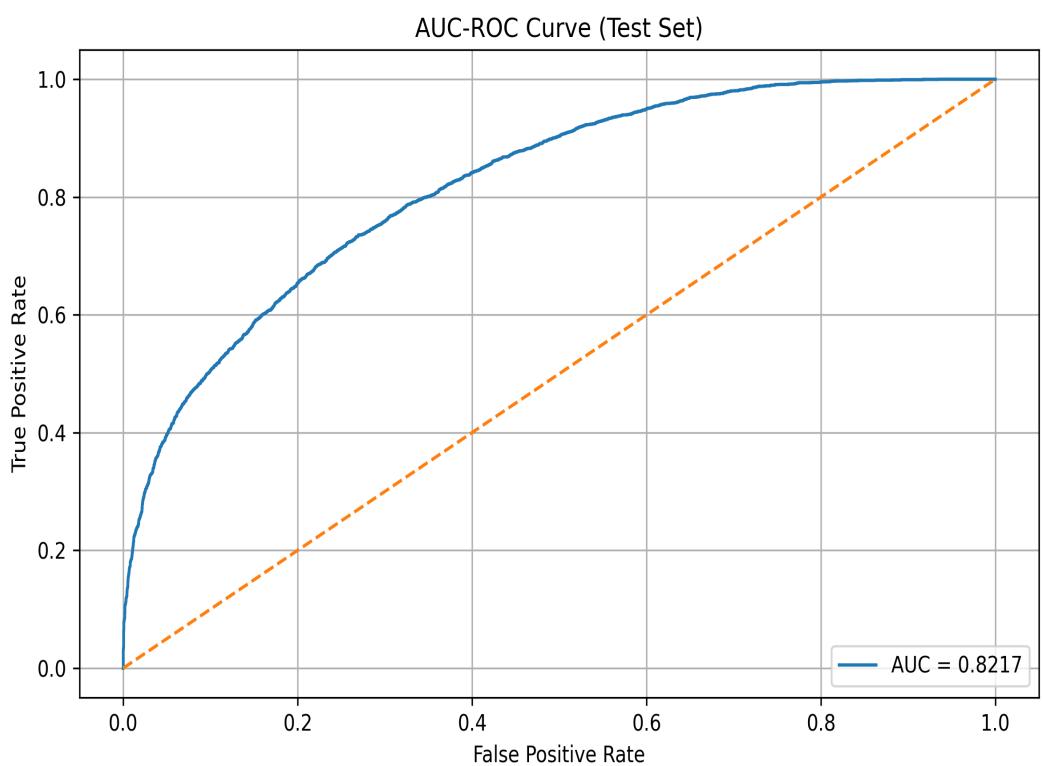
Train vs Validation Loss



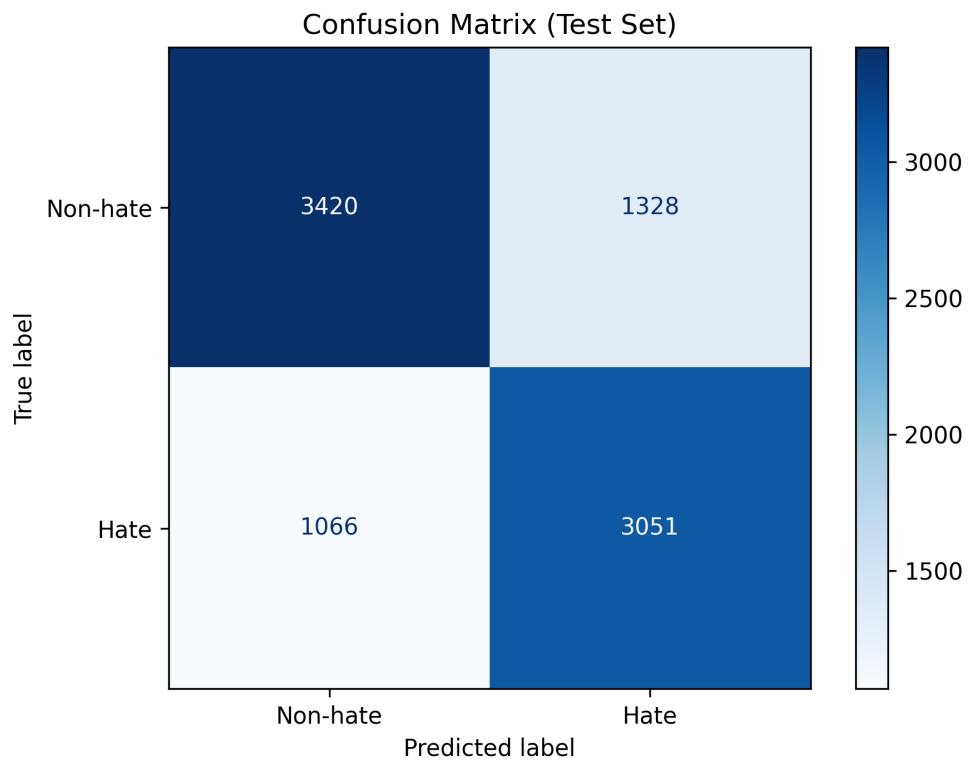
Train vs Validation Accuracy



AUC-ROC Curve



Confusion Matrix



Note: Best model selected using validation F1-score with early stopping. Final metrics are reported on the unseen test set.