

Training Metrics

Metric	Value
Final Training Loss	0.3246
Training Samples Per Second	13.99
Training Loss	0.4245
Total Training Steps	6378
Total Training Epochs	3.0

Loss Progression

Epoch	Loss	Gradient Norm	Learning Rate
0.24	0.5796	4.9346	1.84e-05
0.47	0.5163	2.5236	1.68e-05
0.71	0.5031	4.7232	1.53e-05
0.94	0.4881	4.6845	1.37e-05
1.18	0.4445	3.5594	1.22e-05
1.41	0.4356	4.3994	1.06e-05
1.65	0.4308	12.8999	9.03e-06
1.88	0.4247	5.3029	7.46e-06
2.12	0.3799	8.8403	5.89e-06
2.35	0.3278	12.1735	4.32e-06
2.59	0.3335	9.8386	2.76e-06
2.82	0.3246	8.8726	1.18e-06

Dataset Information

Dataset Characteristic	Count
Total Dataset Size	42,513 examples
Training Set Size	34,010 examples
Test Set Size	8,503 examples
Training : Test Ratio	80% : 20%
Number of Categories	2 (mapped to 5 standard triage levels)

Model Configuration

Parameter	Value
Base Model	emilyalsentzer/Bio_ClinicalBERT
Number of Parameters	~110M
Maximum Sequence Length	128 tokens
Training Batch Size	16
Number of Epochs	3
Initial Learning Rate	2E-05
Weight Decay	0.01
Optimizer	AdamW
Training Hardware	Apple MPS (Metal Performance Shaders)

Inference Performance

Metric	Value
Inference Time Per Sample	~0.1 seconds (on CPU)
Model Size	~450MB