

Model Training Documentation

March 10, 2025

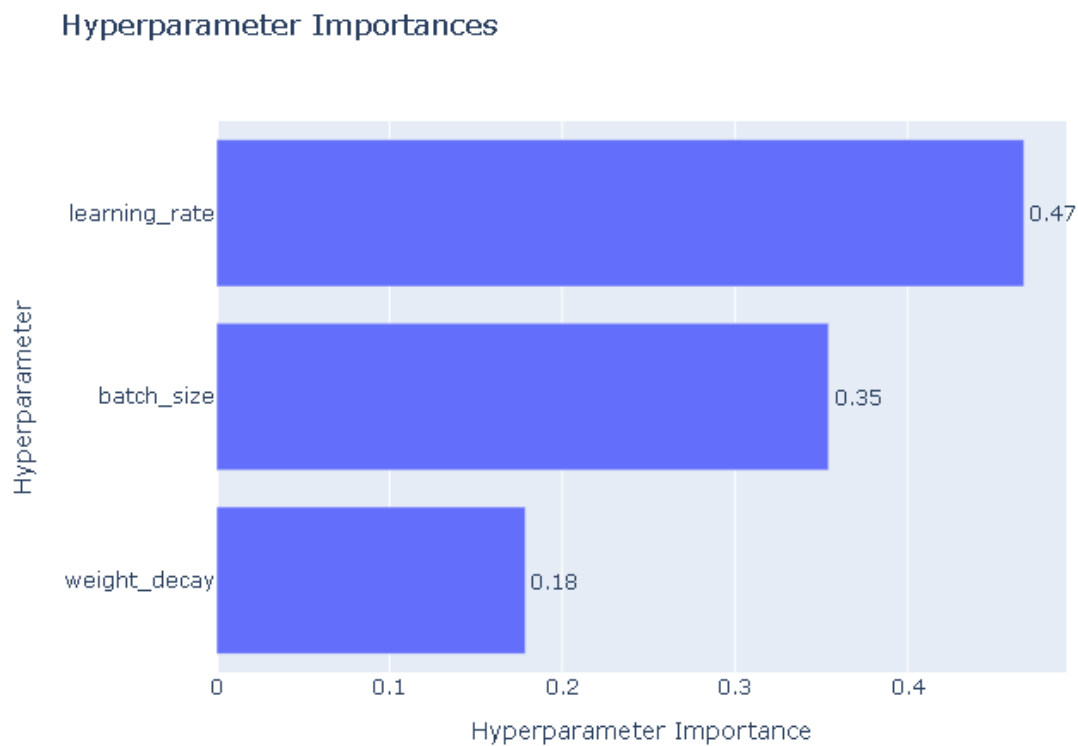
1 Models Comparision

Model Name	Accuracy	F1	Precision	Recall
Roberta	0.8438	0.8378	0.8443	0.8438
ALBERT	0.8327	0.8225	0.8315	0.8327
BertBase	0.8358	0.8253	0.8320	0.8358
DistilBERT	0.8223	0.8102	0.8126	0.8223
ELECTRA	0.8259	0.8124	0.8244	0.8259
XLNET	0.8288	0.8207	0.8298	0.8288

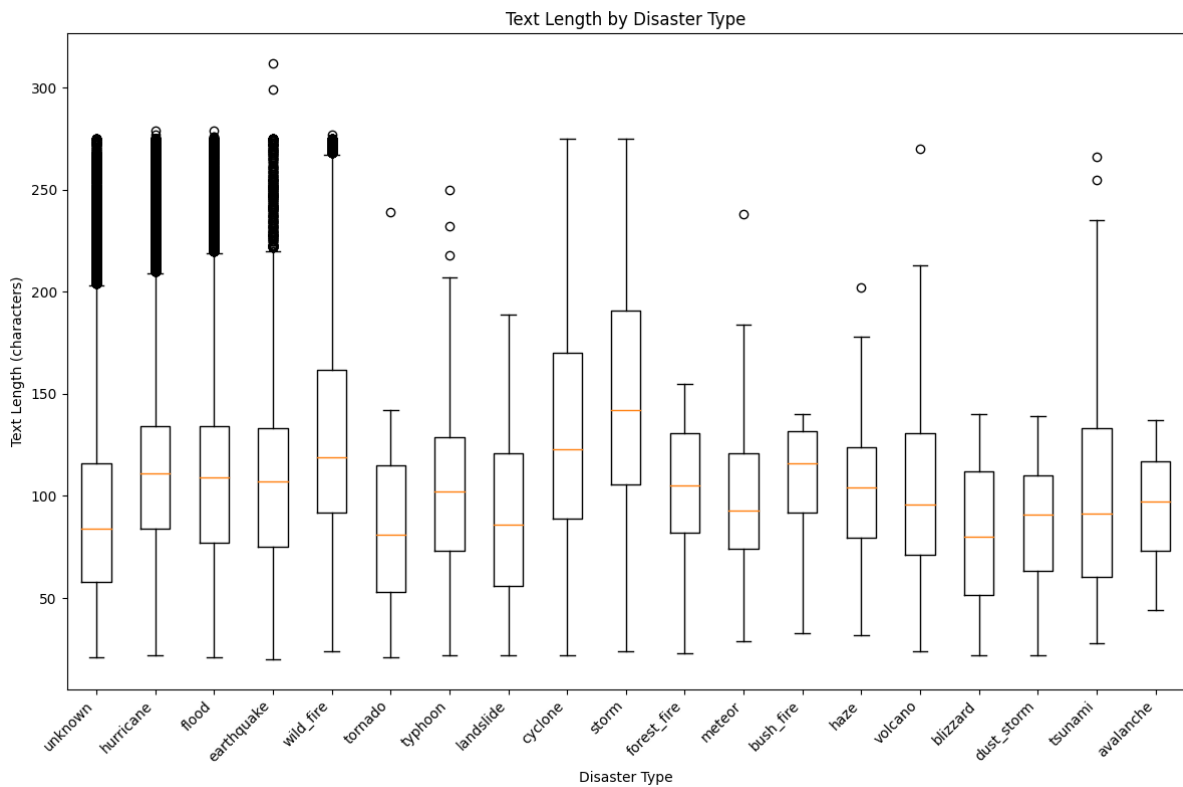
Table 1: Comparison of Model Performance

2 Output (RoBERTa)

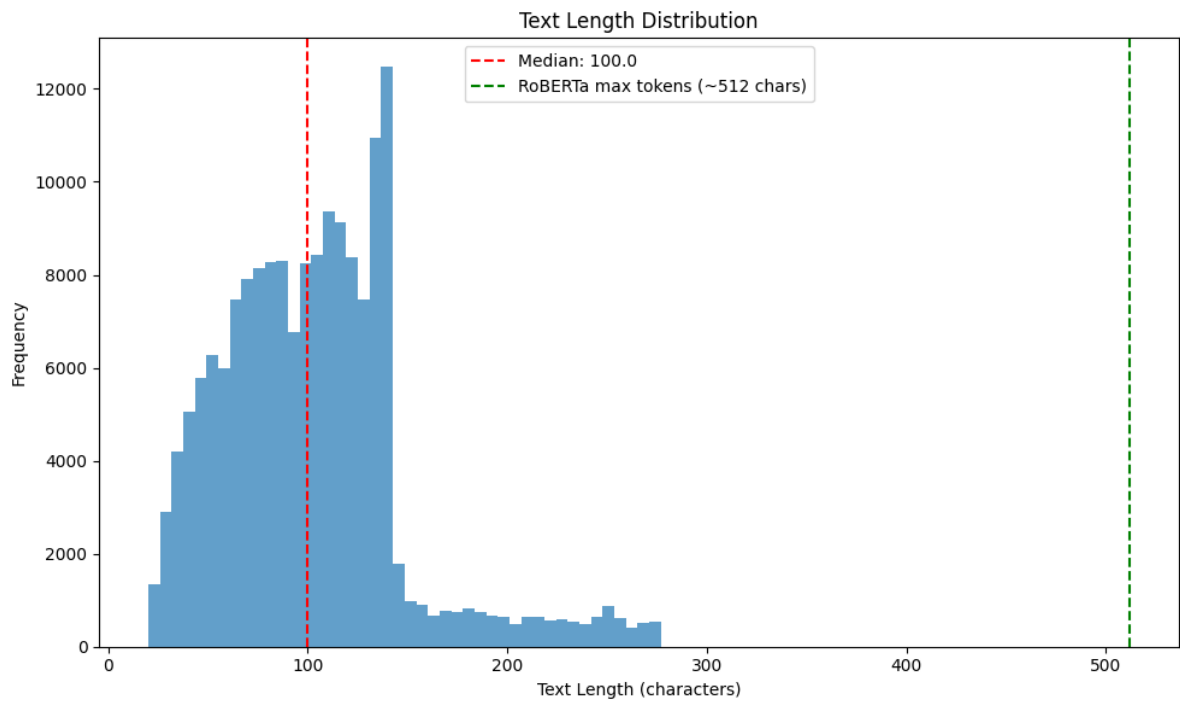
2.1 Hyperparameter Importances



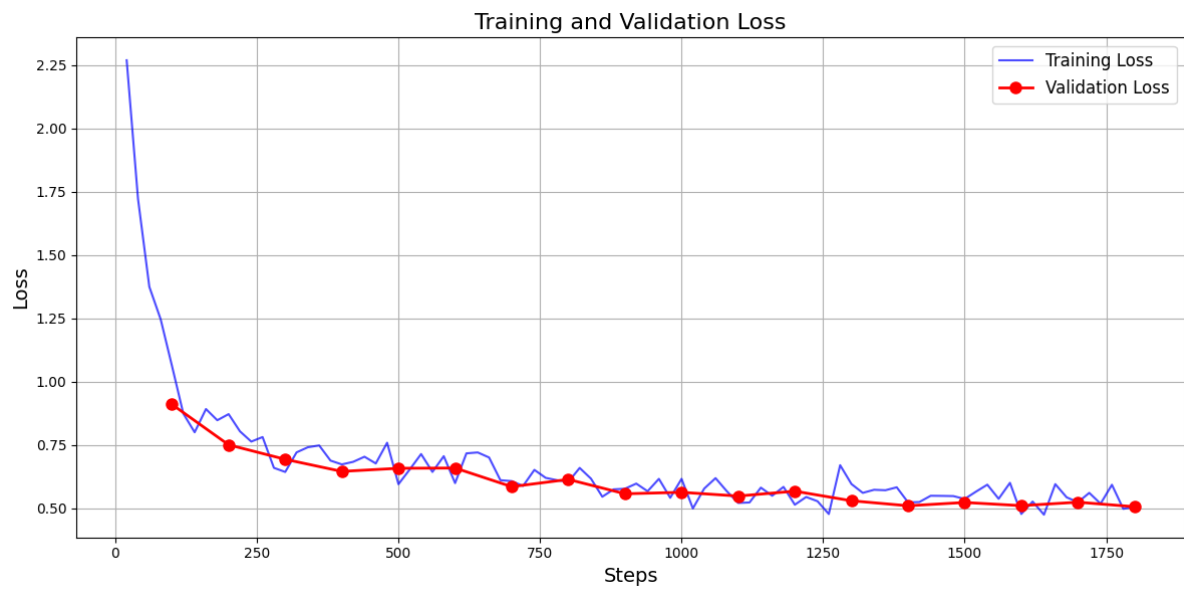
2.2 Text Length by Disaster Type



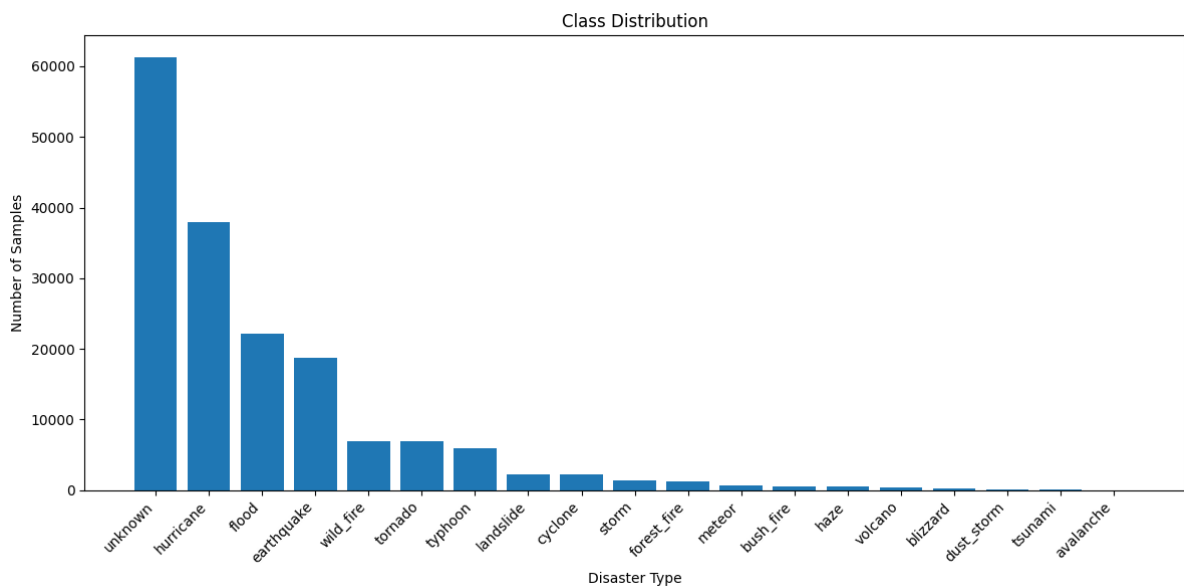
2.3 Text Length Distribution



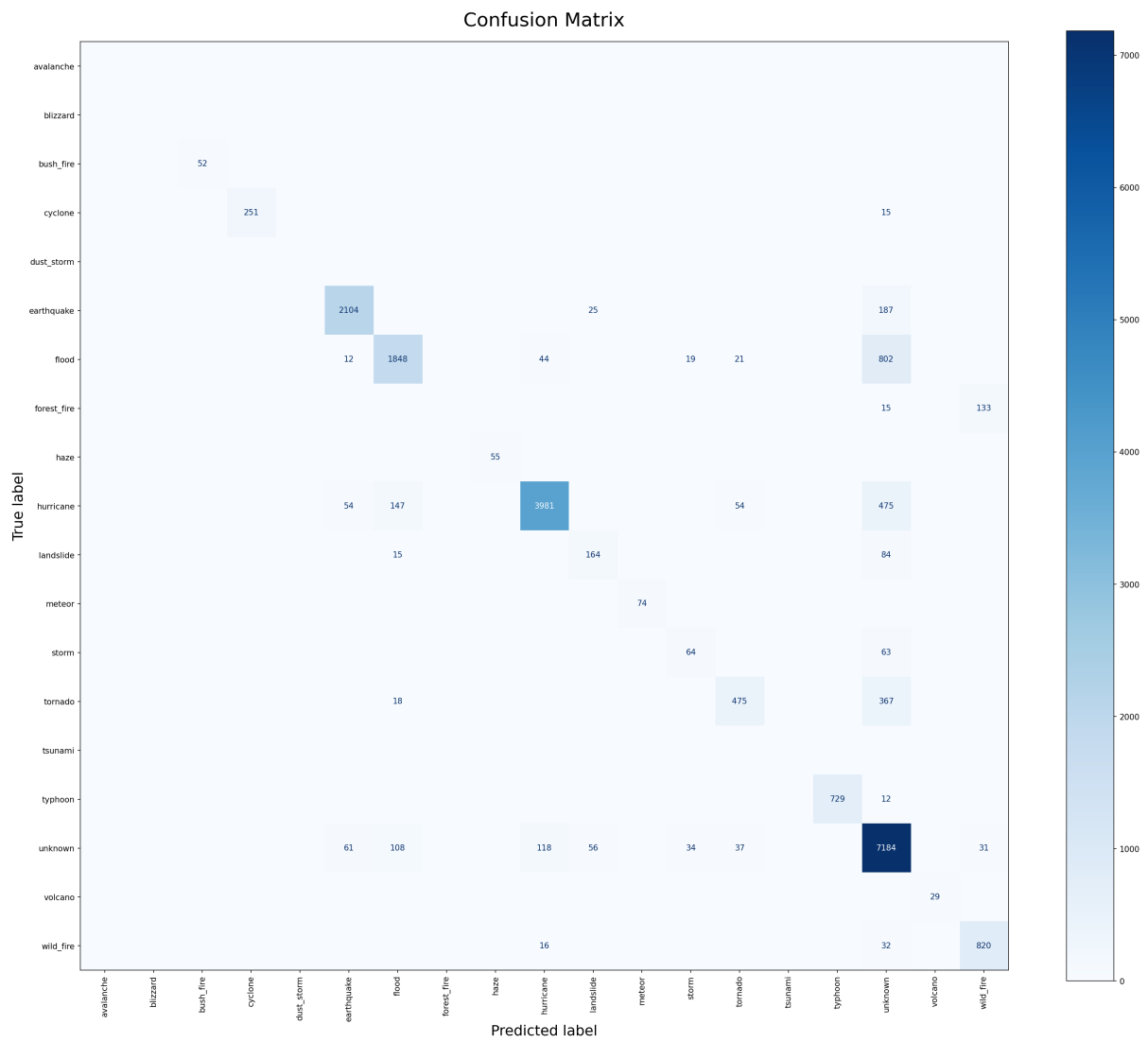
2.4 Training and Validation Loss



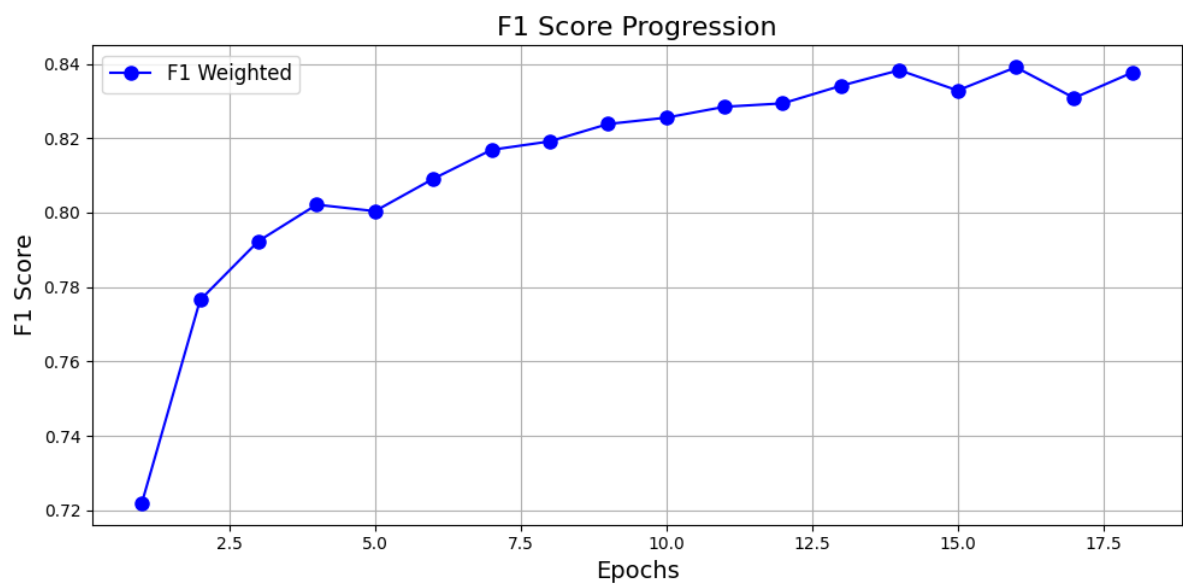
2.5 Class Distribution



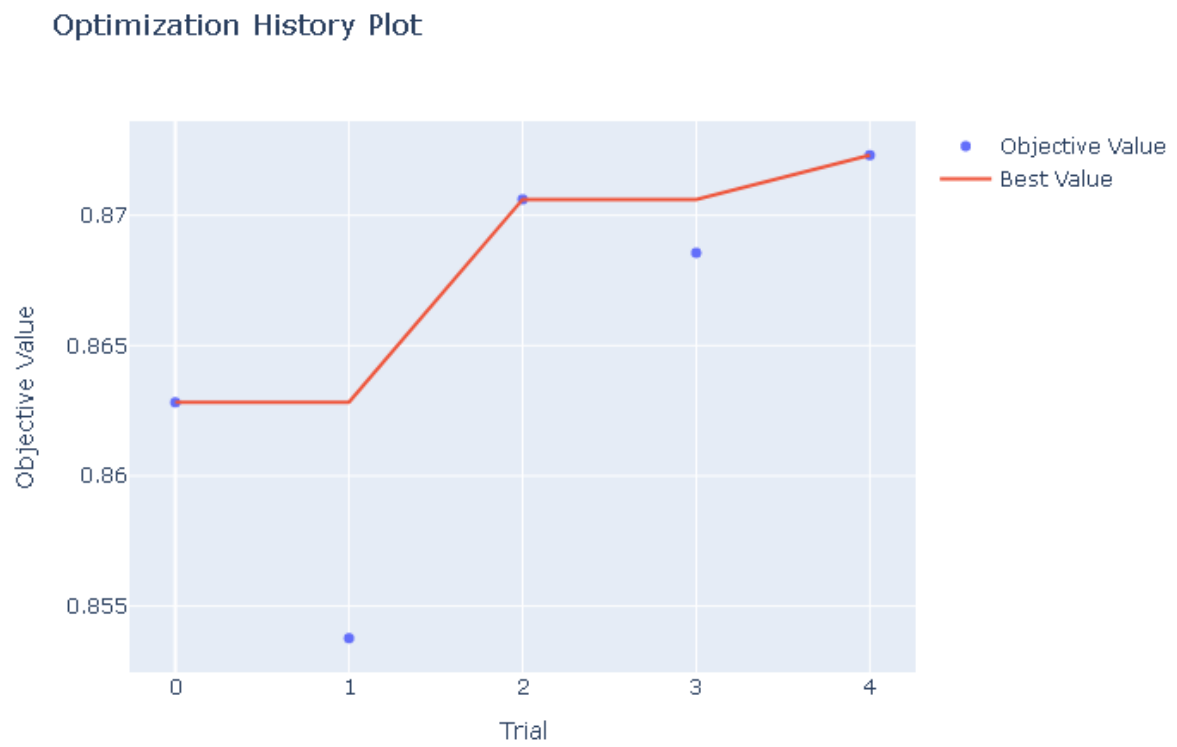
2.6 Confusion Matrix



2.7 F1 Score Progression

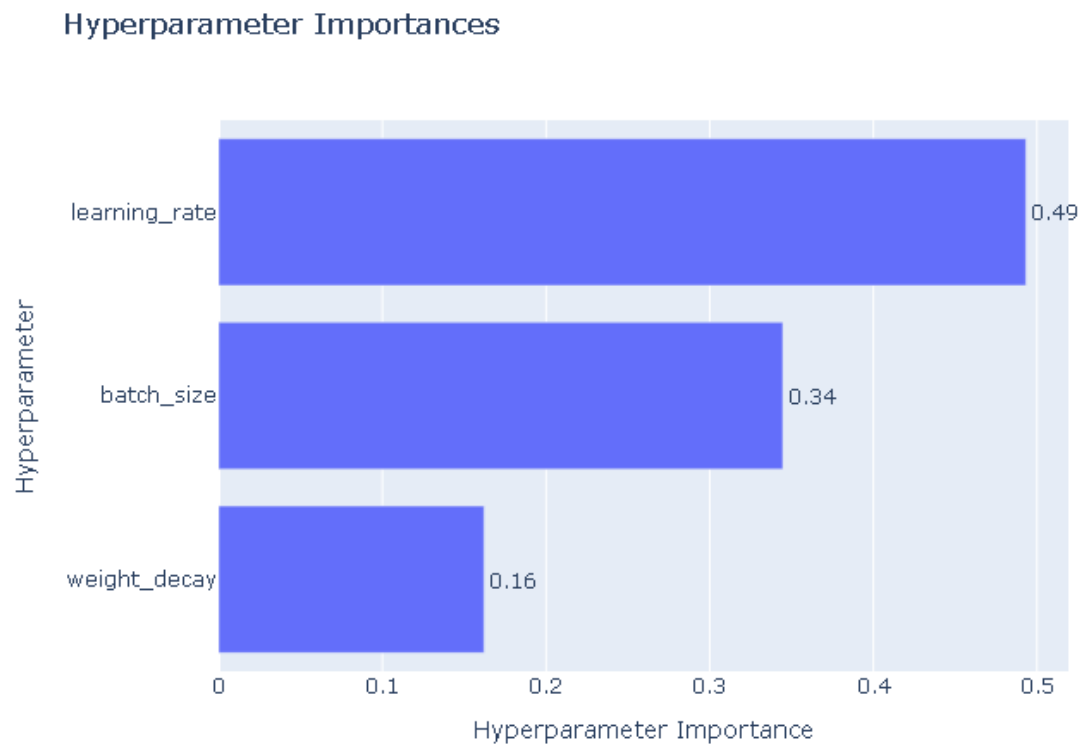


2.8 Optimization History Plot

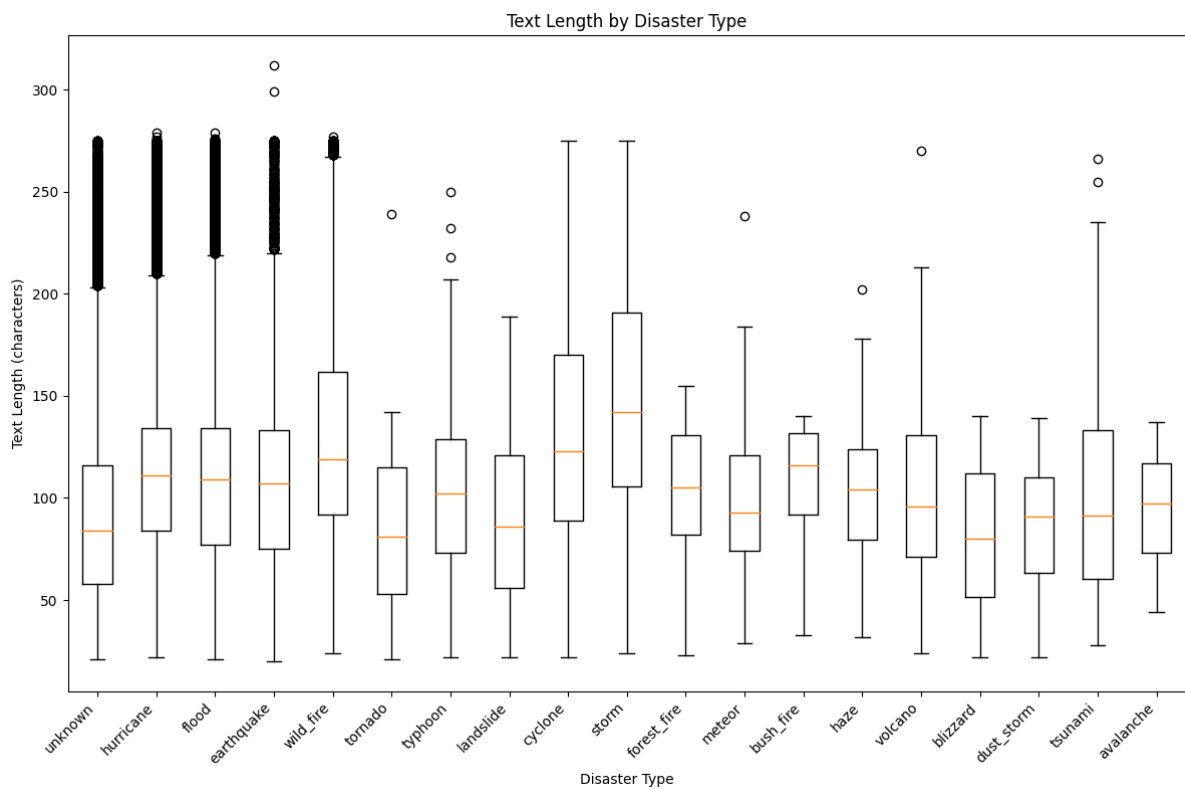


3 Output (ALBERT)

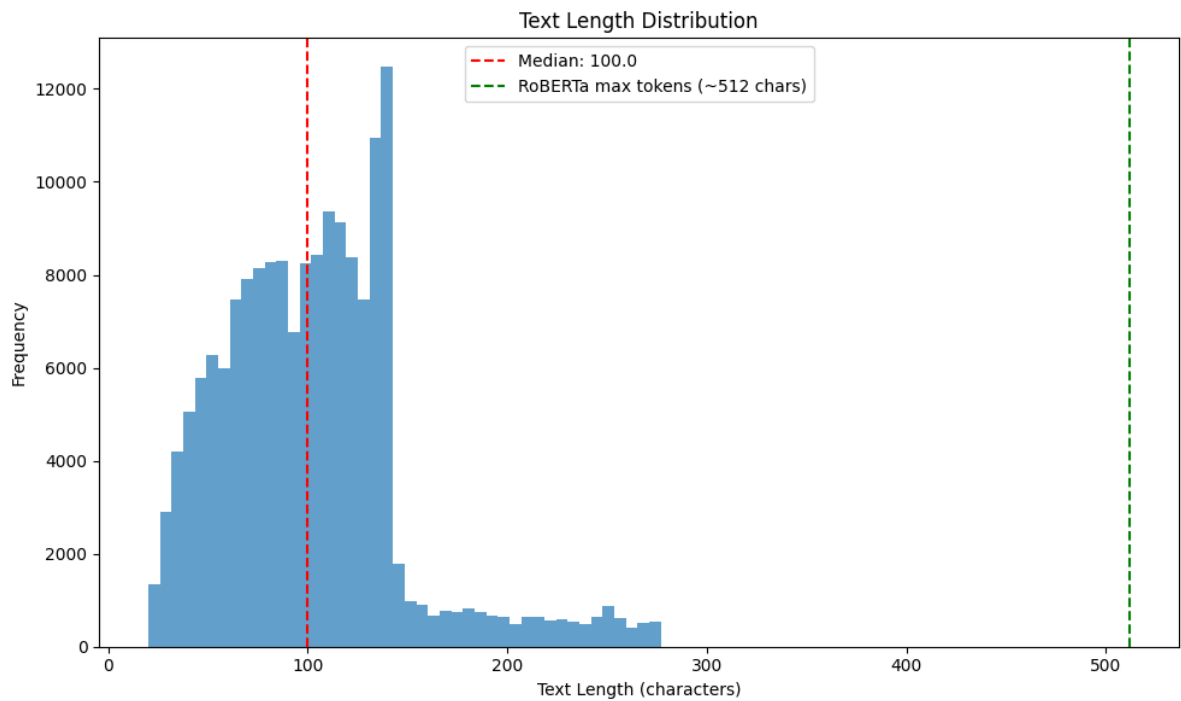
3.1 Hyperparameter Importances



3.2 Text Length by Disaster Type



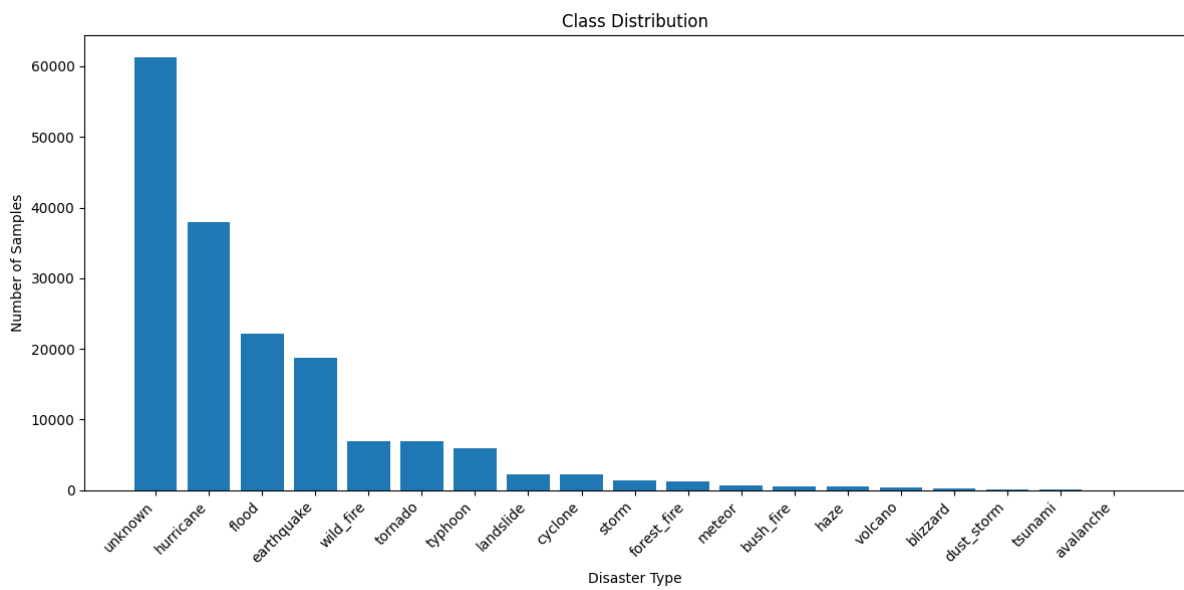
3.3 Text Length Distribution



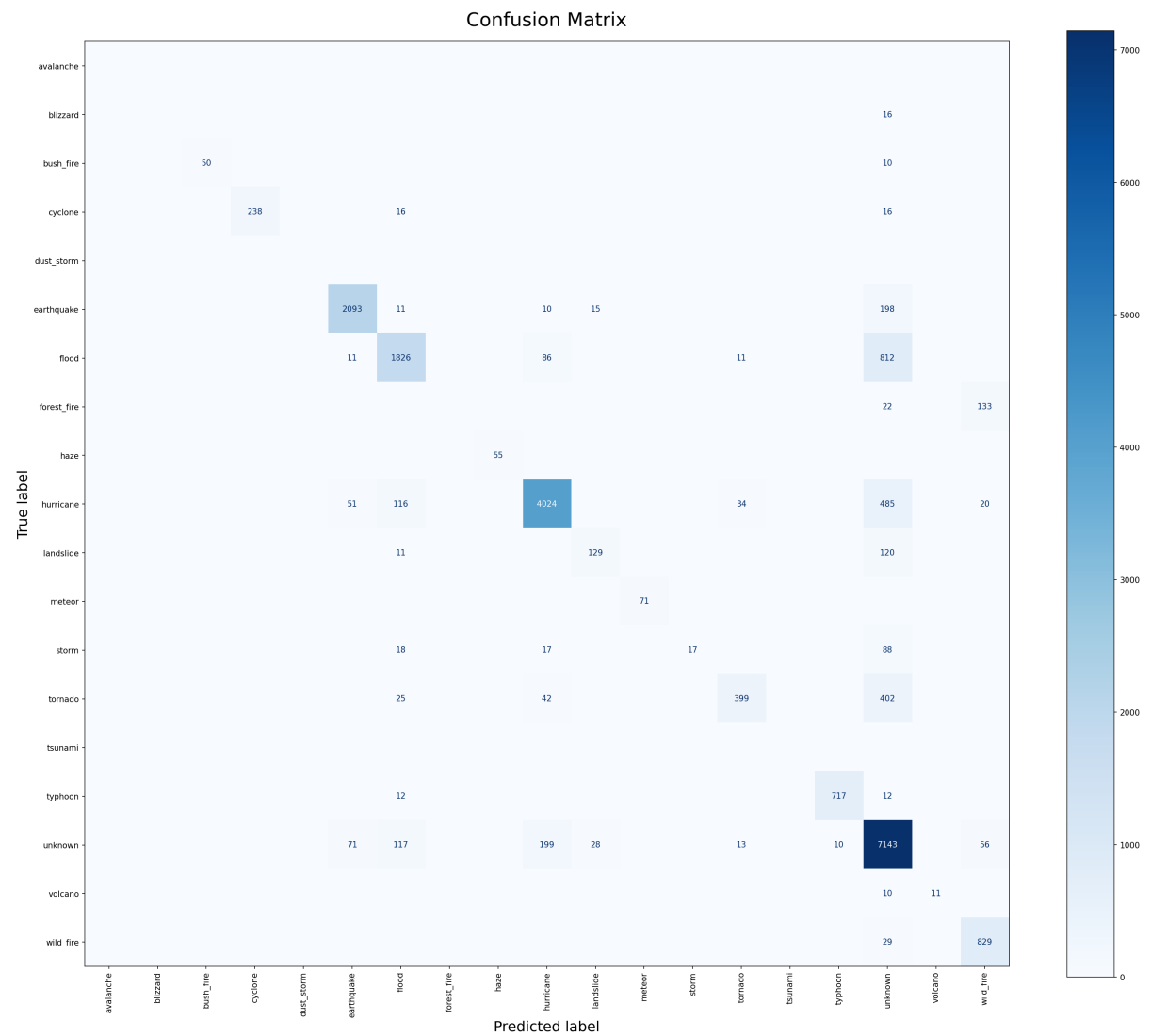
3.4 Training and Validation Loss



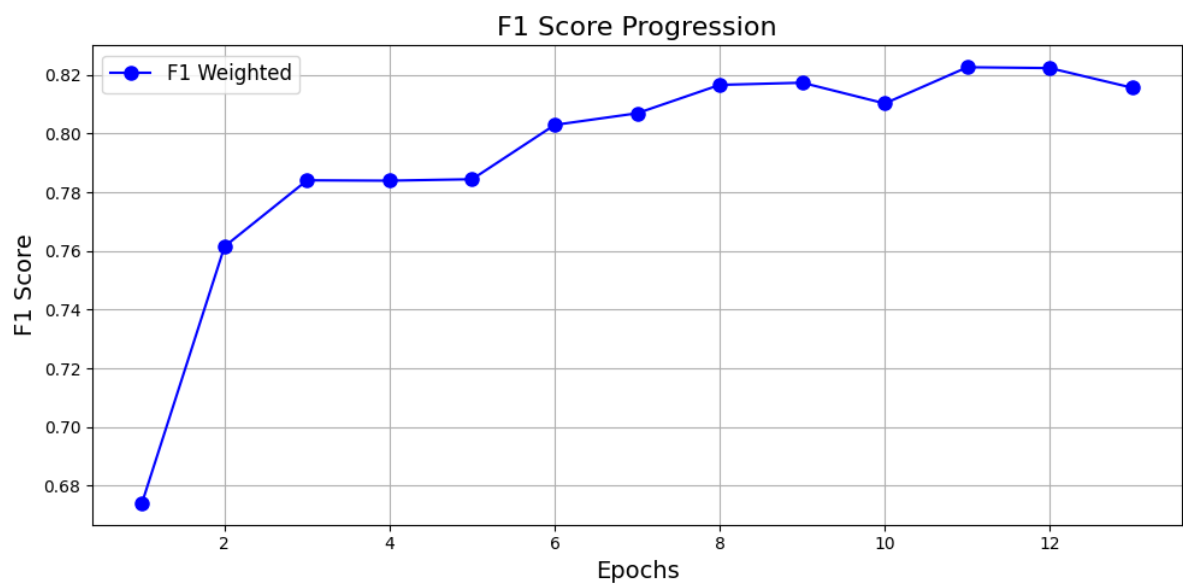
3.5 Class Distribution



3.6 Confusion Matrix

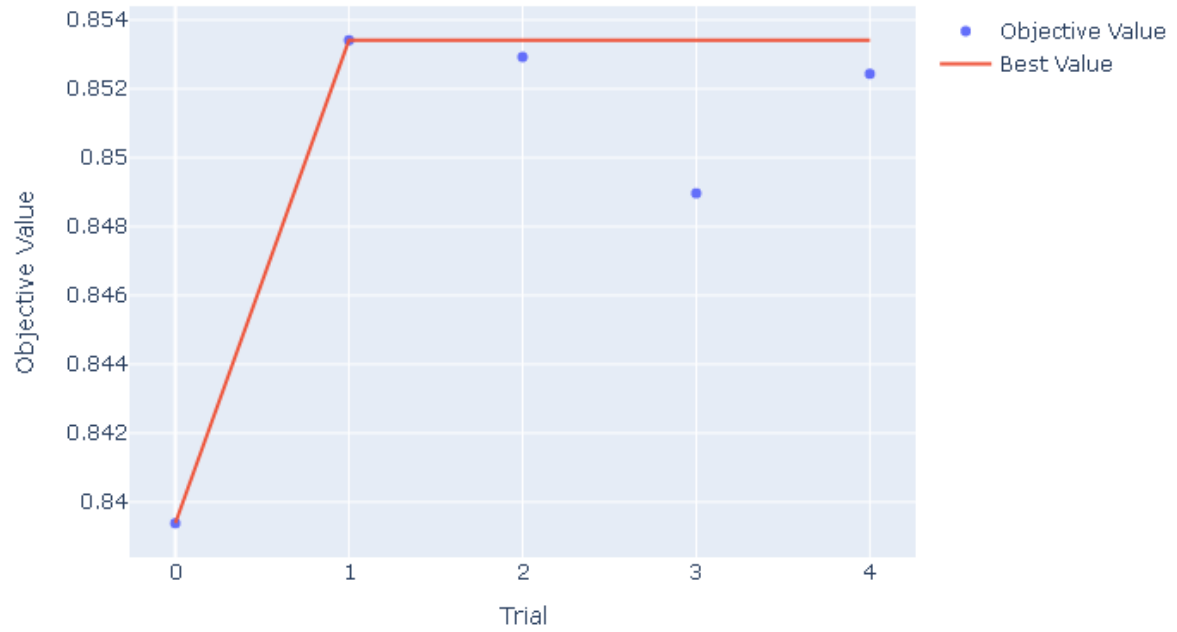


3.7 F1 Score Progression



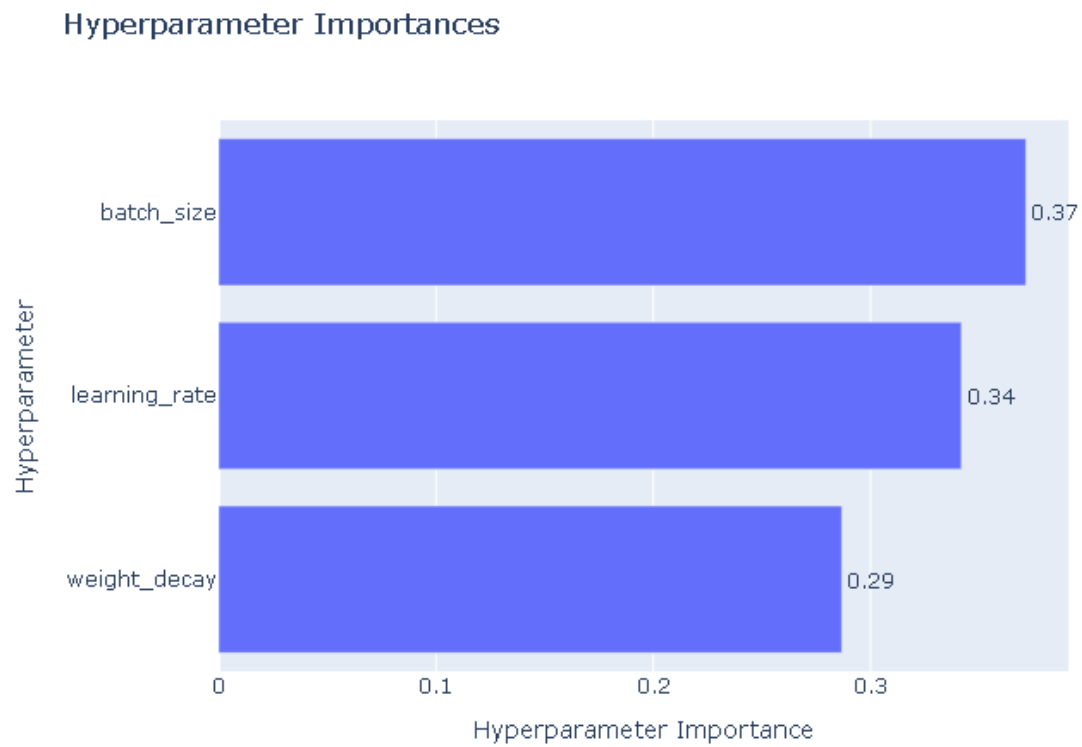
3.8 Optimization History Plot

Optimization History Plot

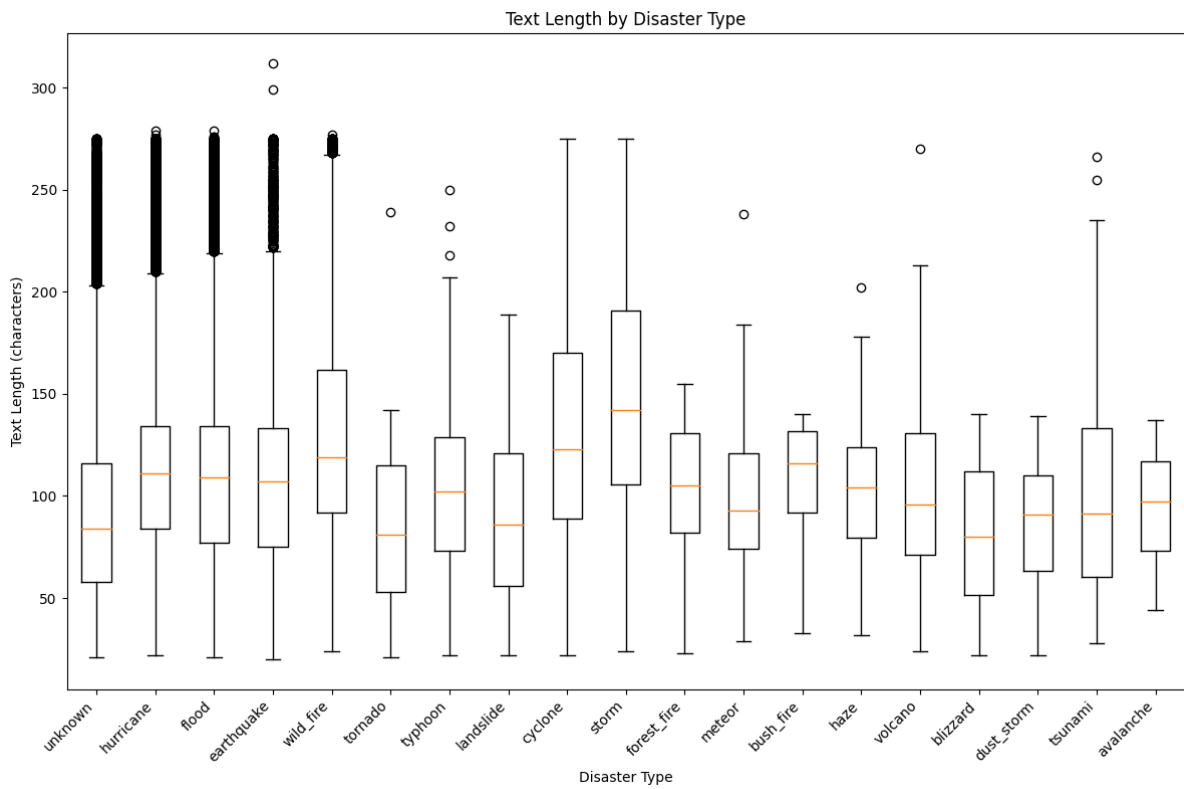


4 Output (BertBase)

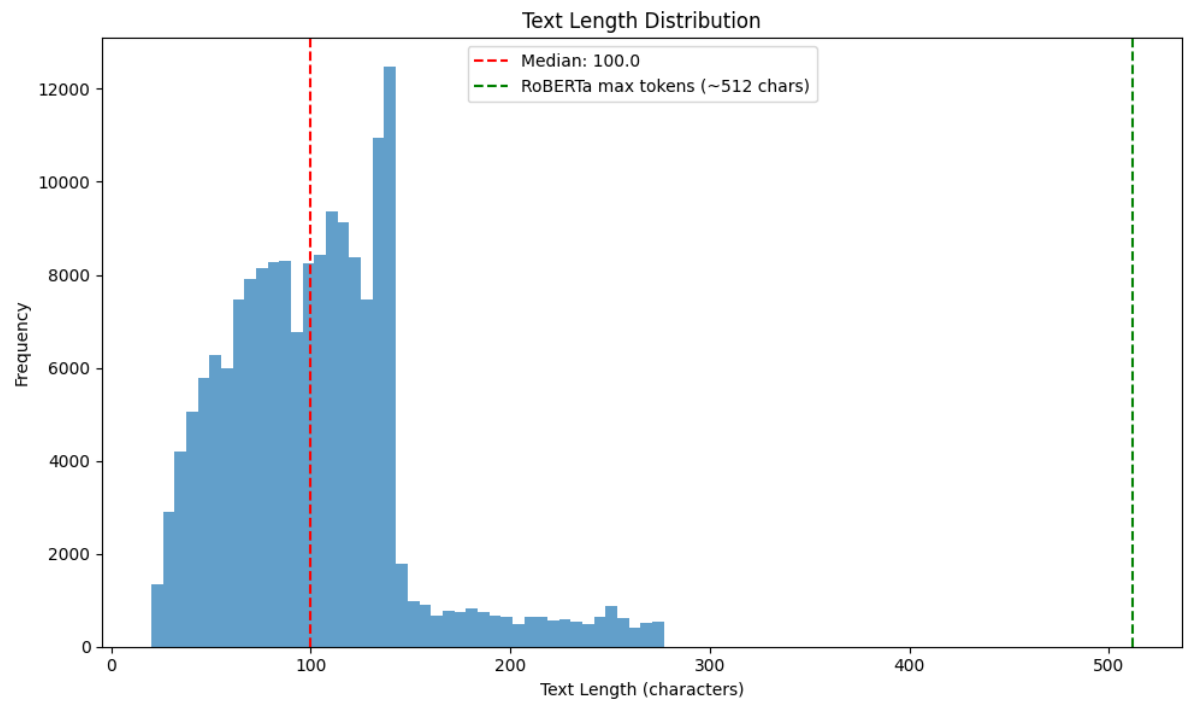
4.1 Hyperparameter Importances



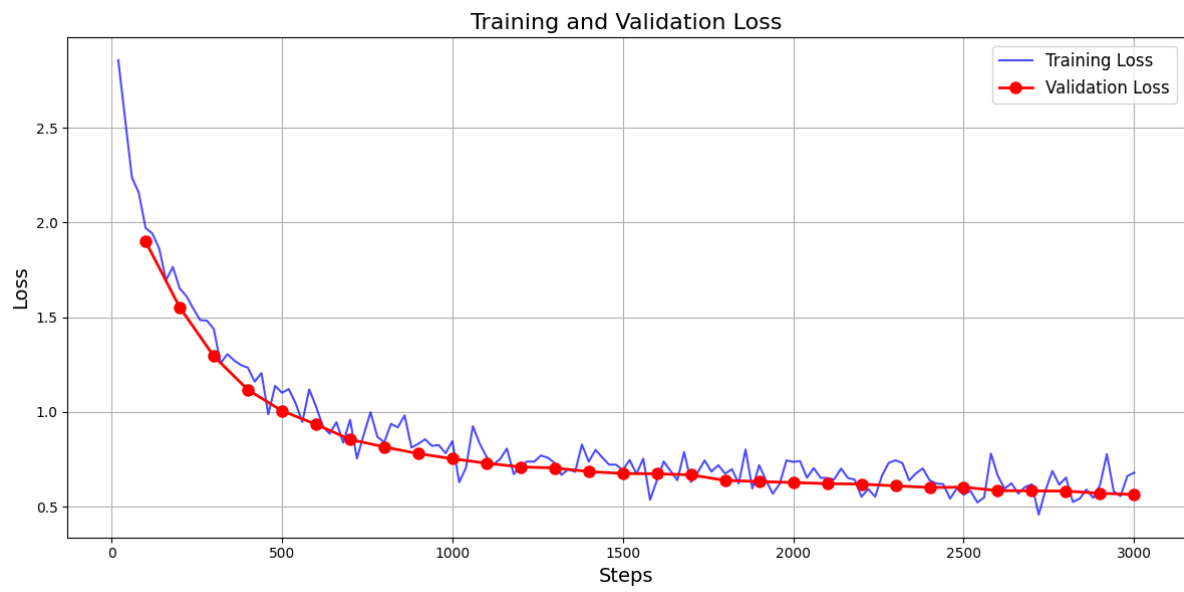
4.2 Text Length by Disaster Type



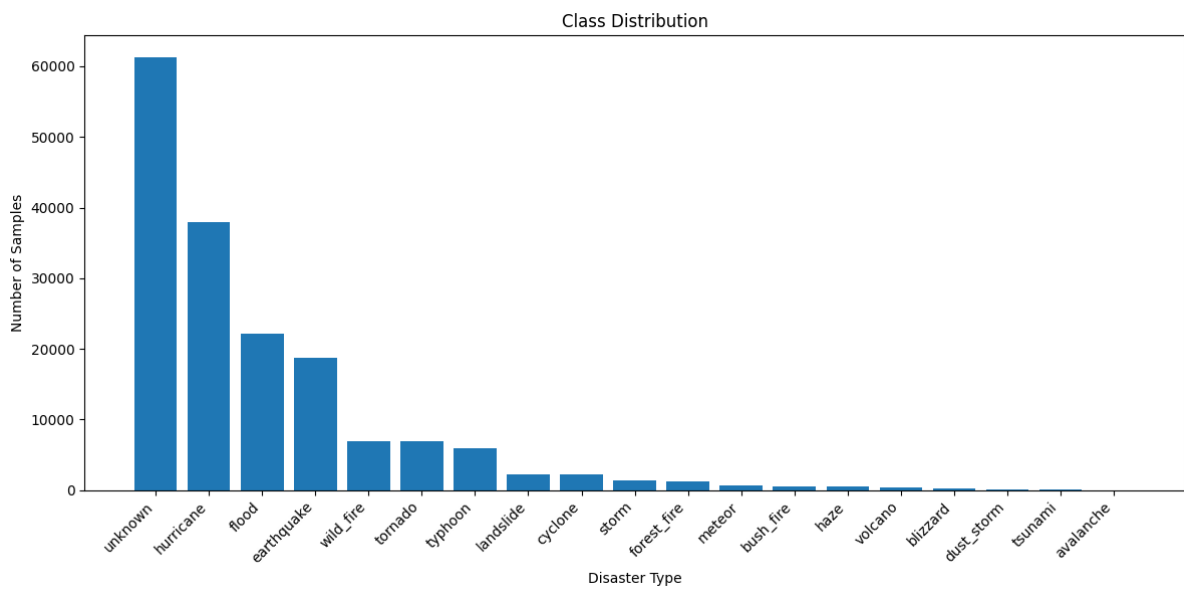
4.3 Text Length Distribution



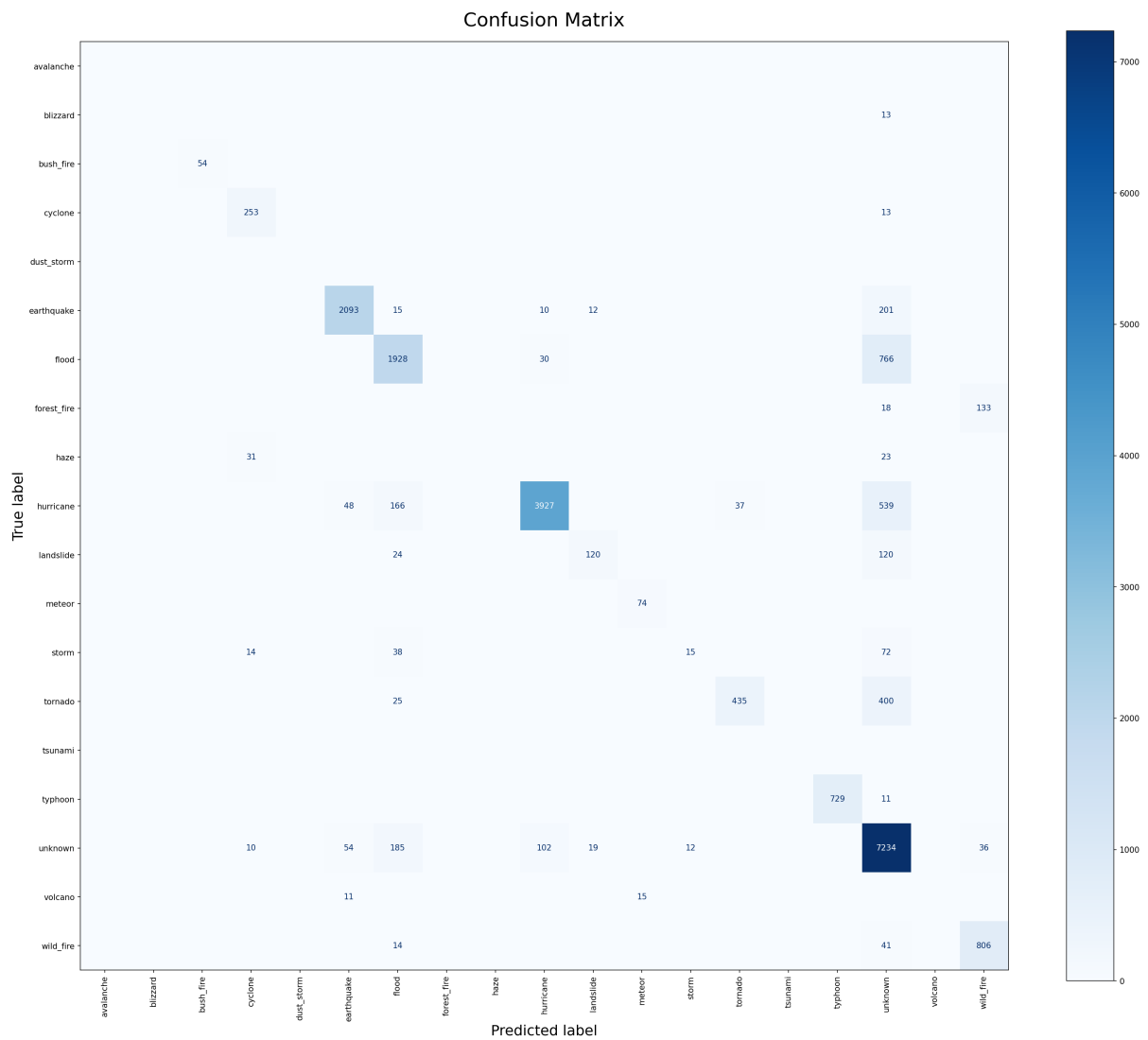
4.4 Training and Validation Loss



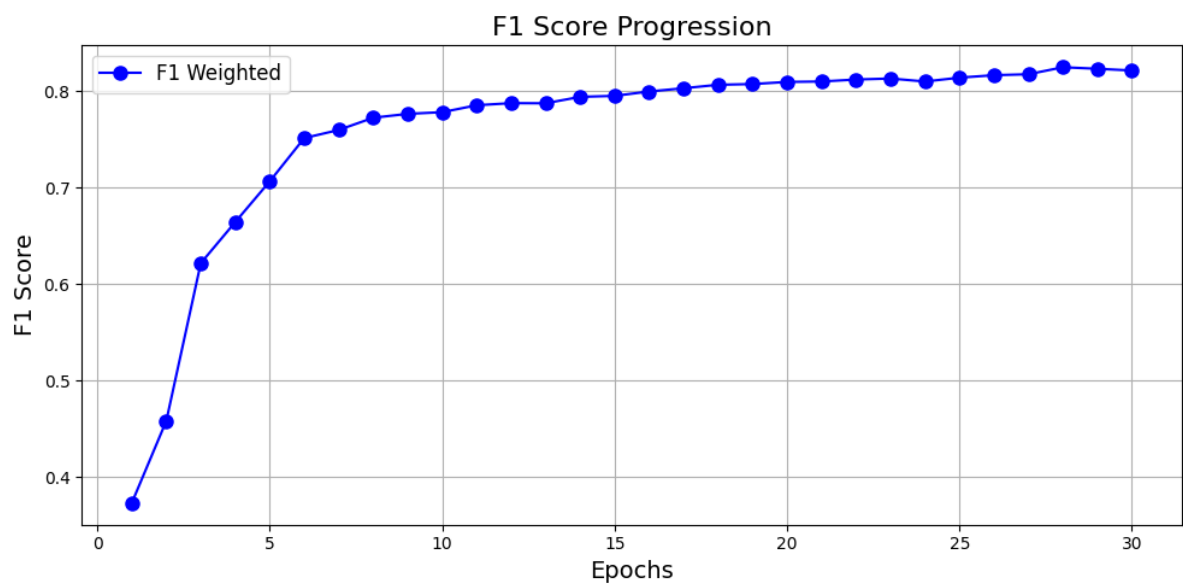
4.5 Class Distribution



4.6 Confusion Matrix

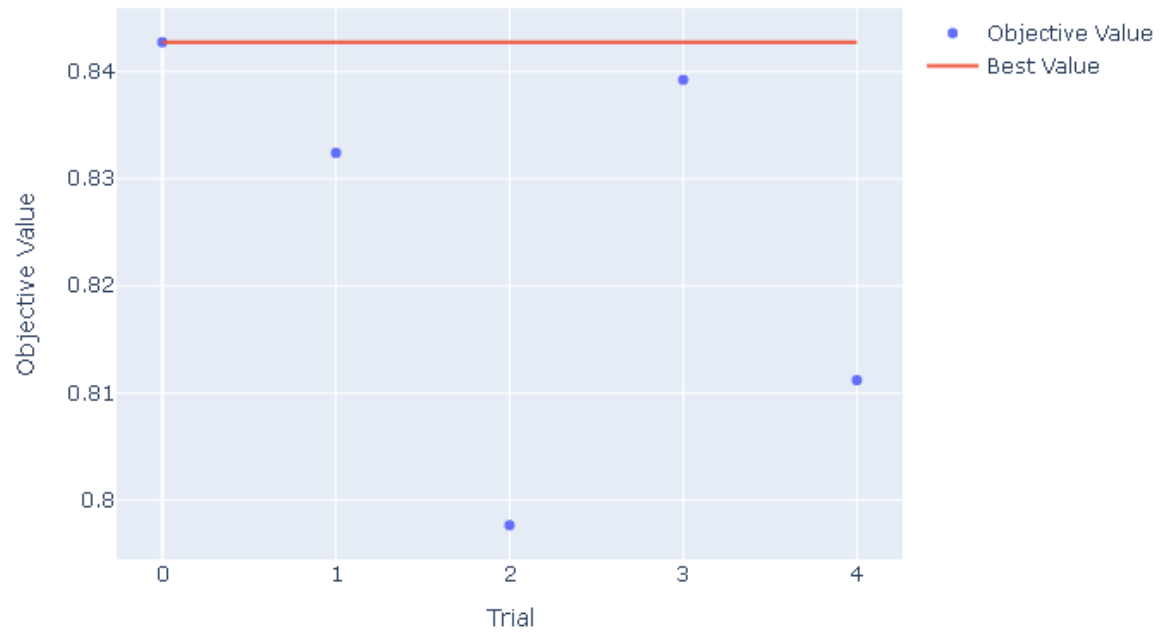


4.7 F1 Score Progression



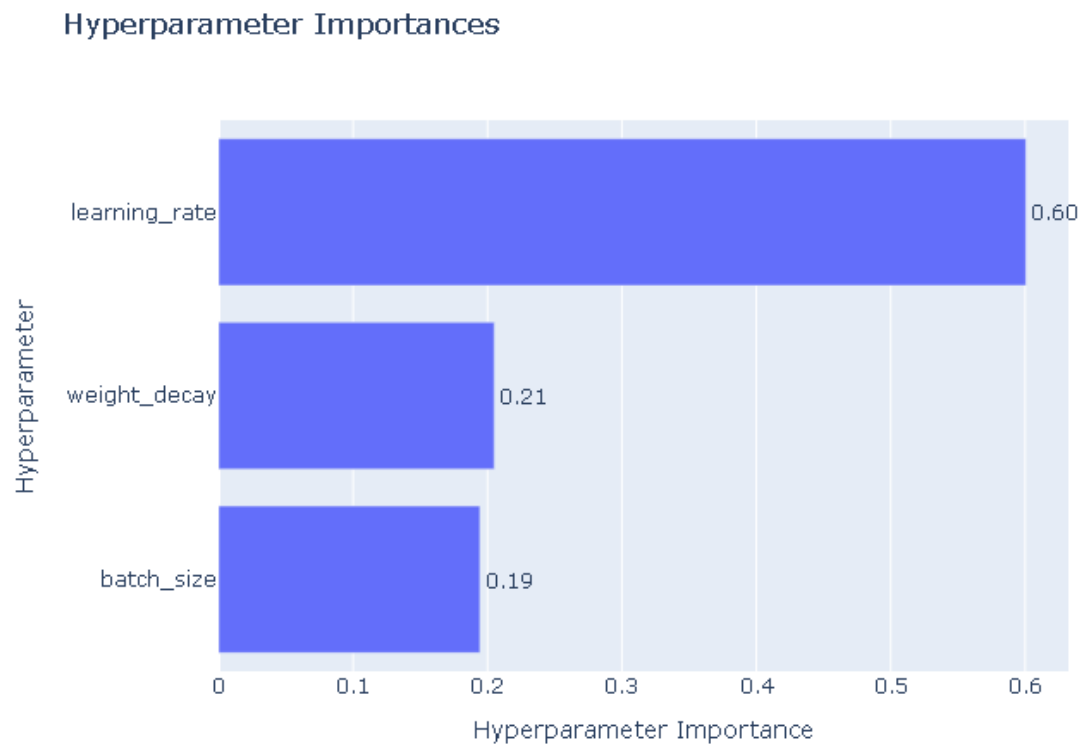
4.8 Optimization History Plot

Optimization History Plot

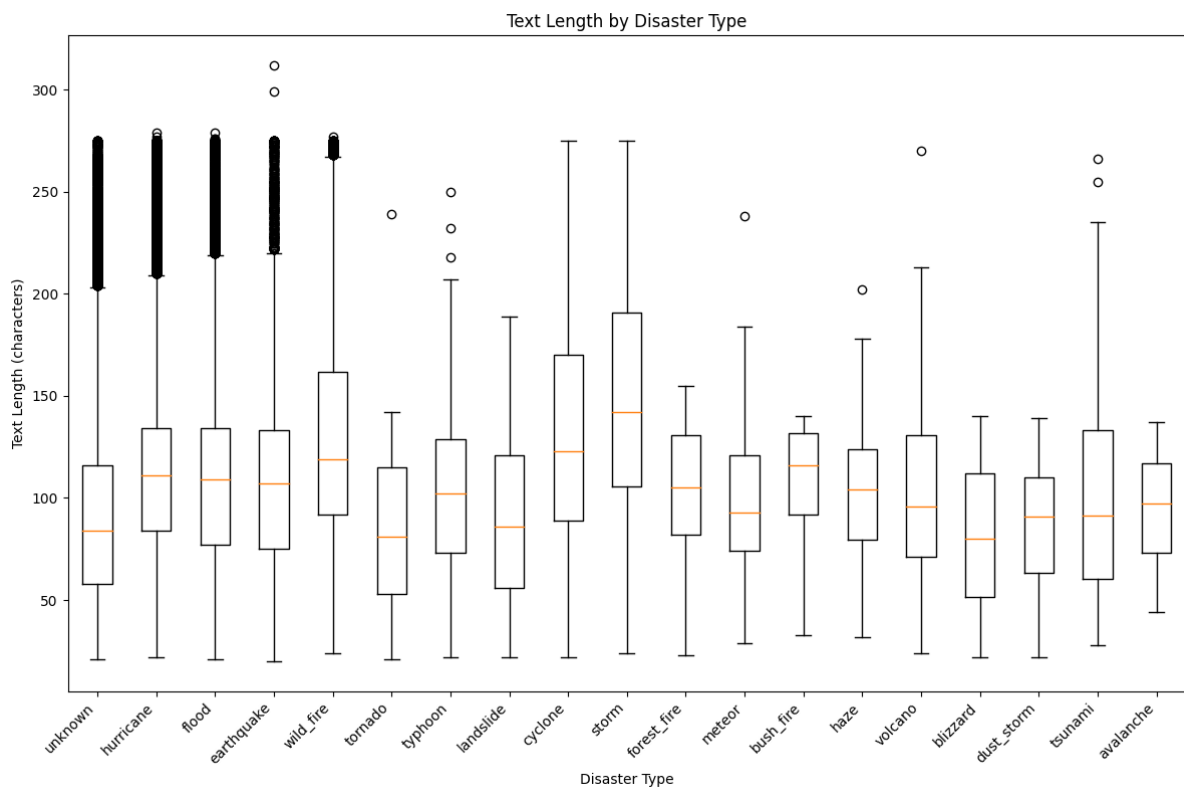


5 Output (distilBERT)

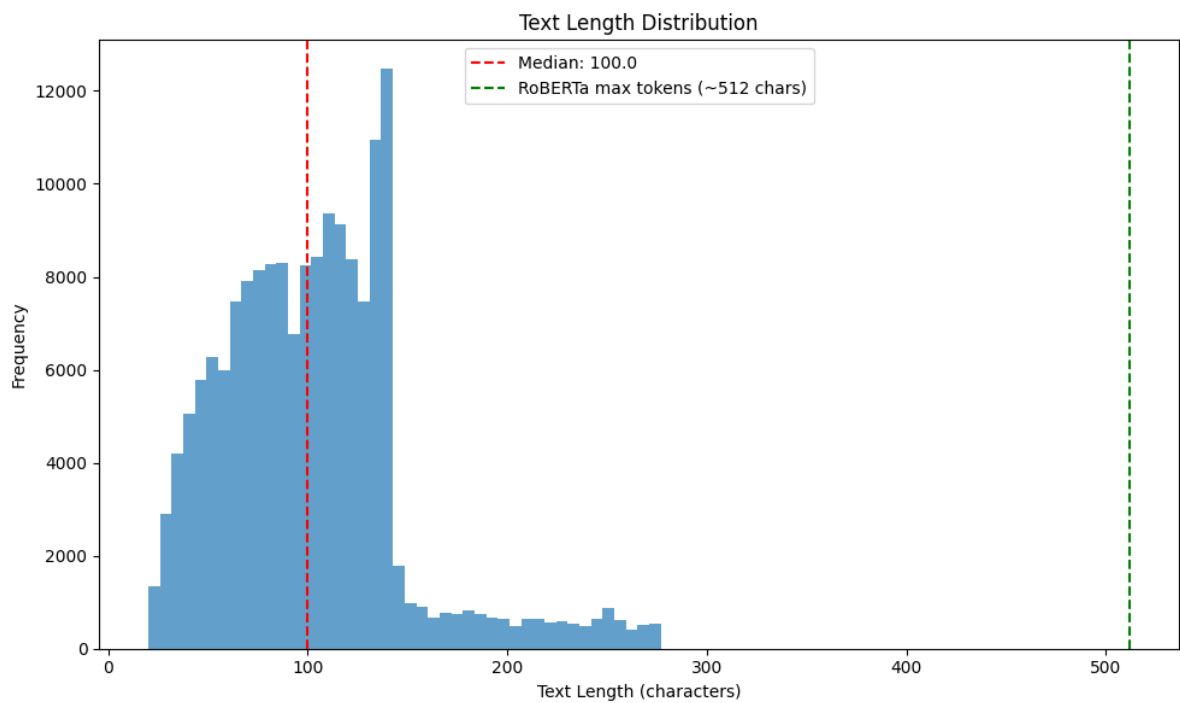
5.1 Hyperparameter Importances



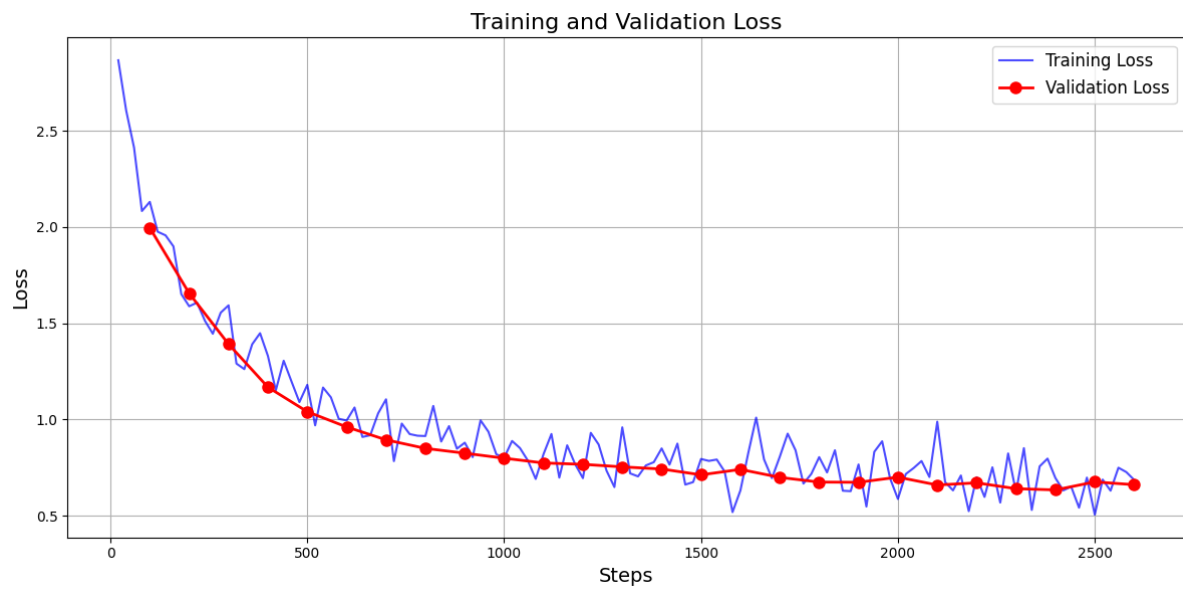
5.2 Text Length by Disaster Type



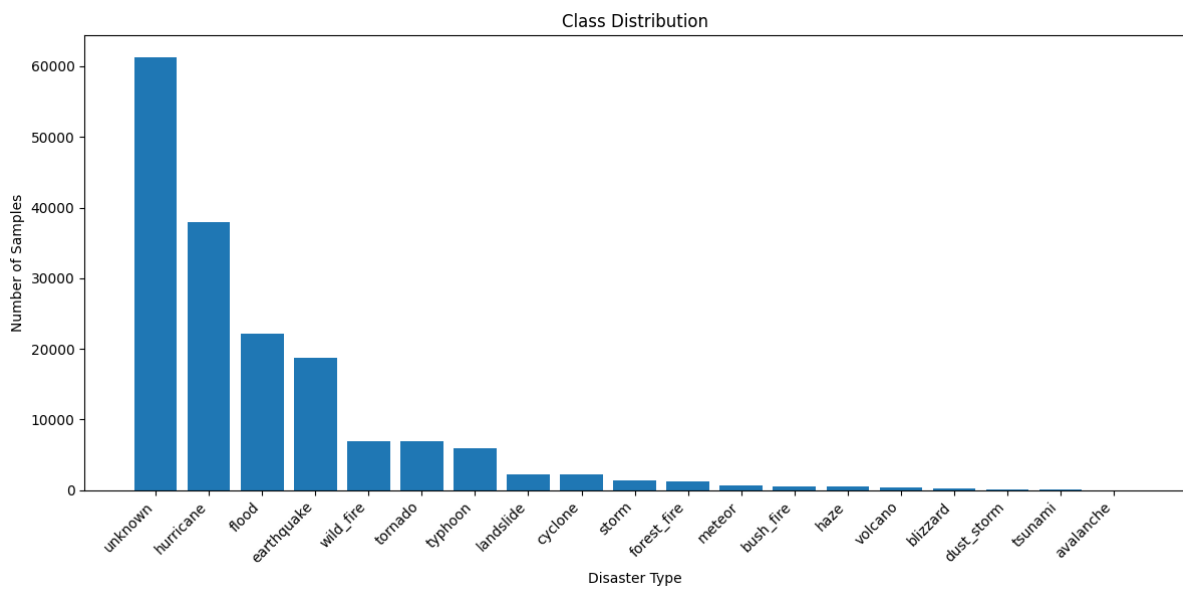
5.3 Text Length Distribution



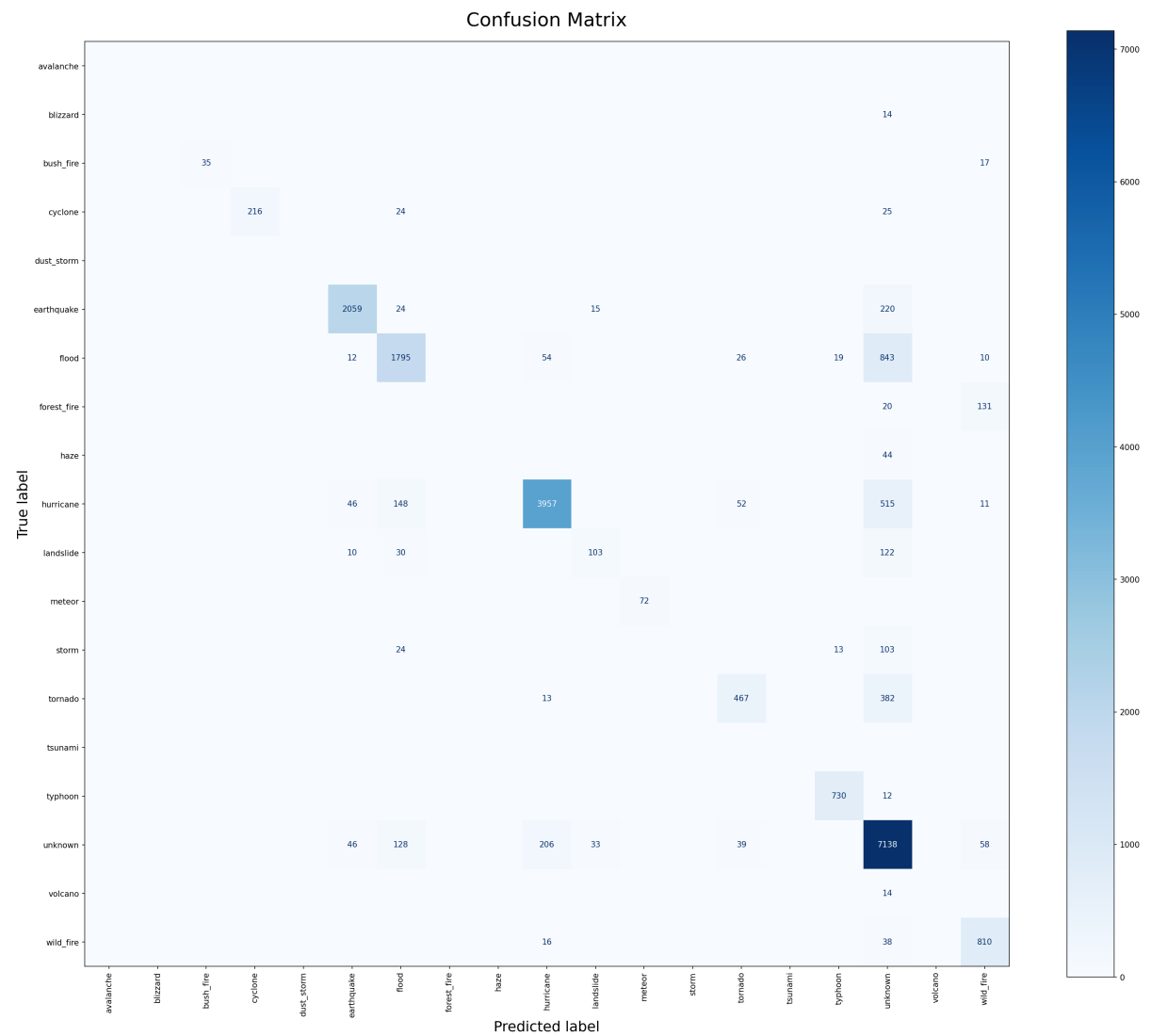
5.4 Training and Validation Loss



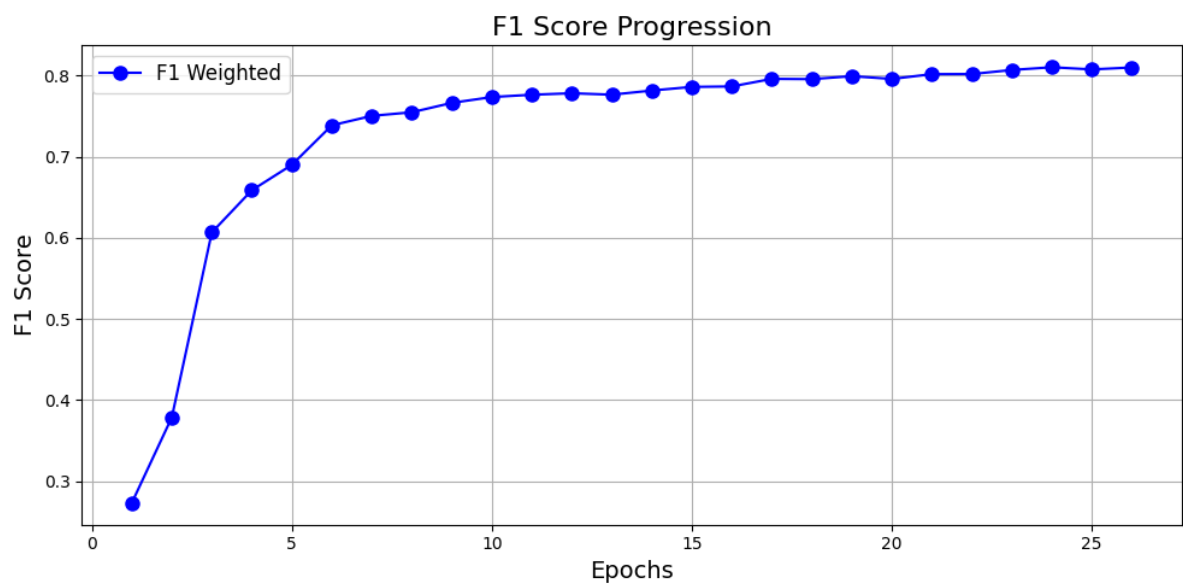
5.5 Class Distribution



5.6 Confusion Matrix

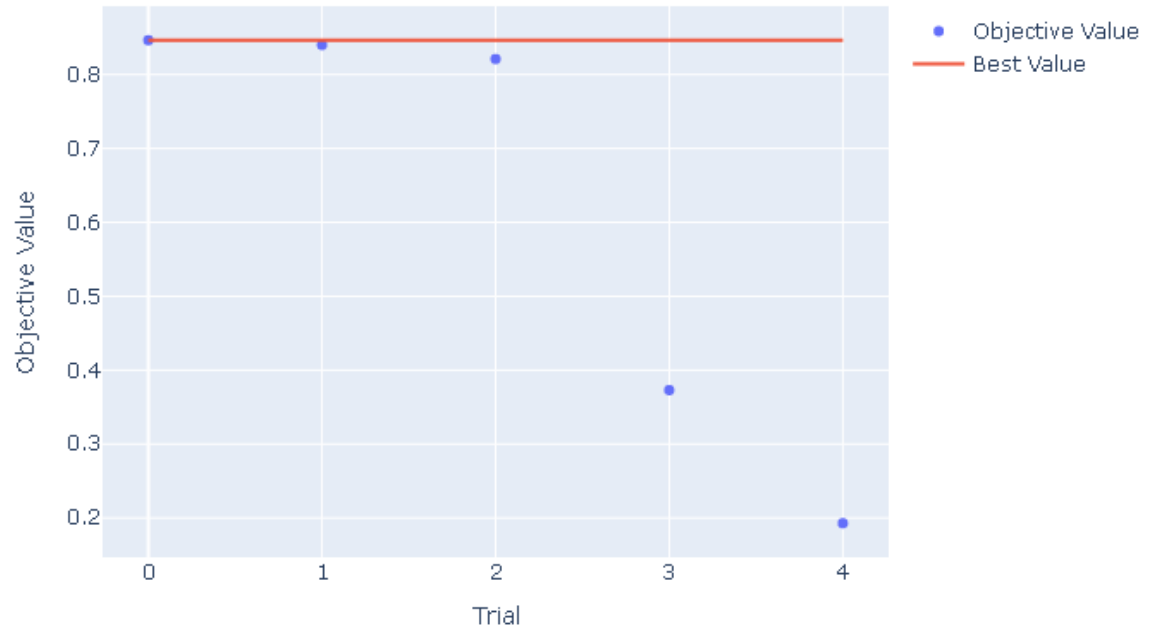


5.7 F1 Score Progression



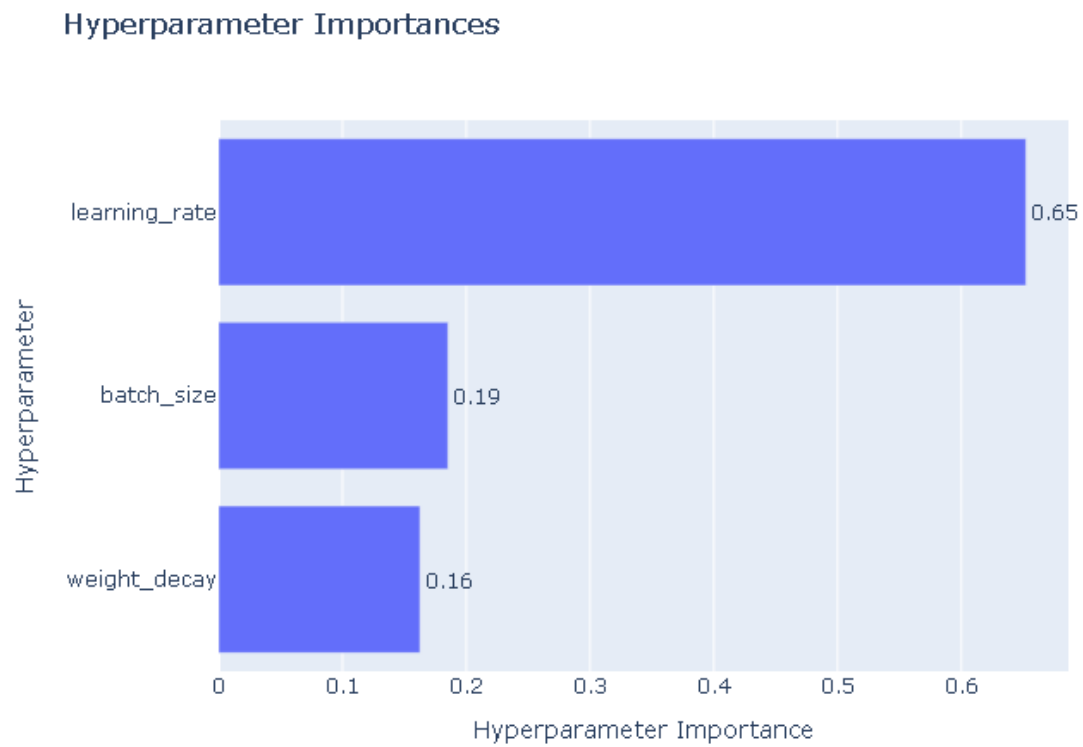
5.8 Optimization History Plot

Optimization History Plot

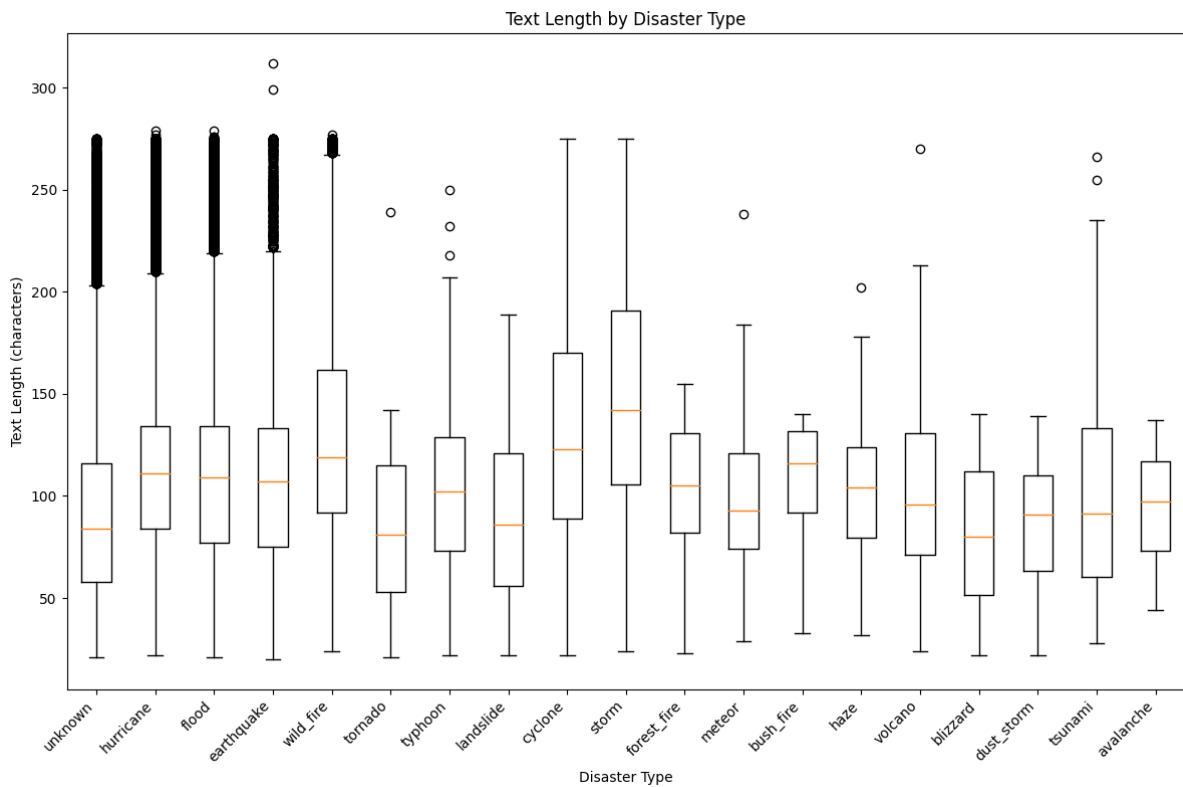


6 Output (ELECTRA)

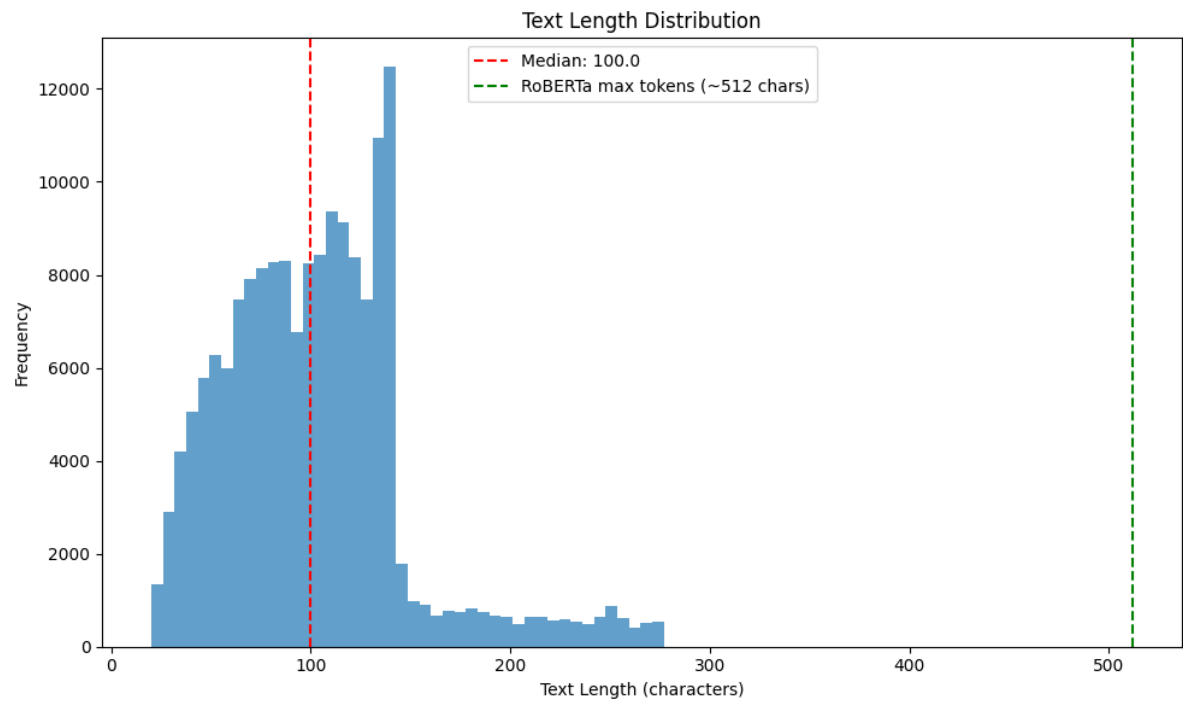
6.1 Hyperparameter Importances



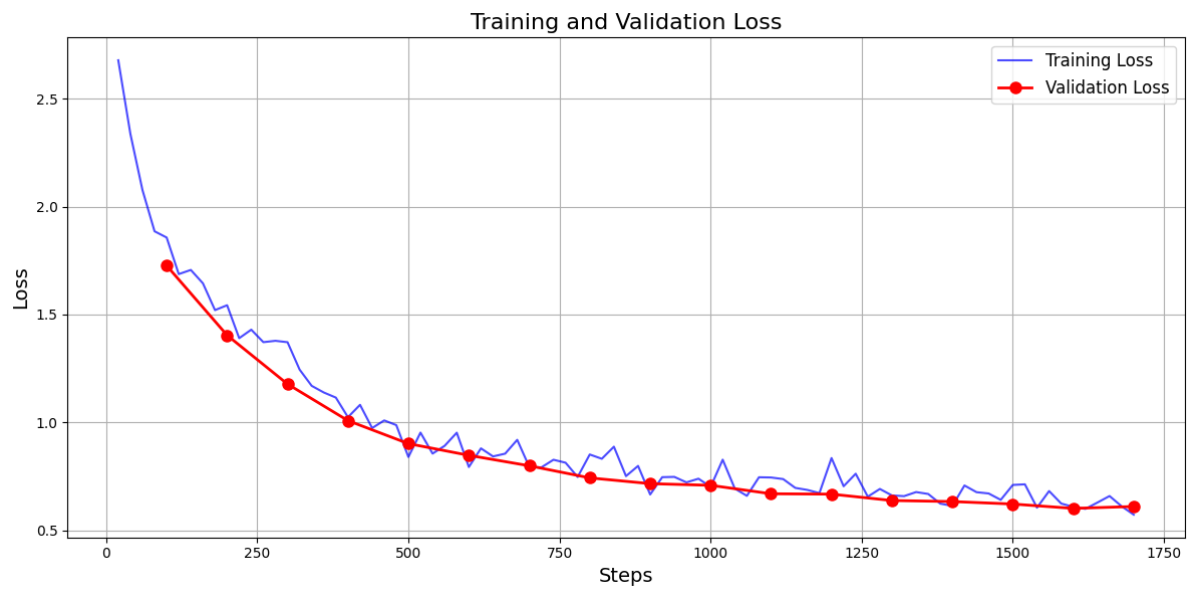
6.2 Text Length by Disaster Type



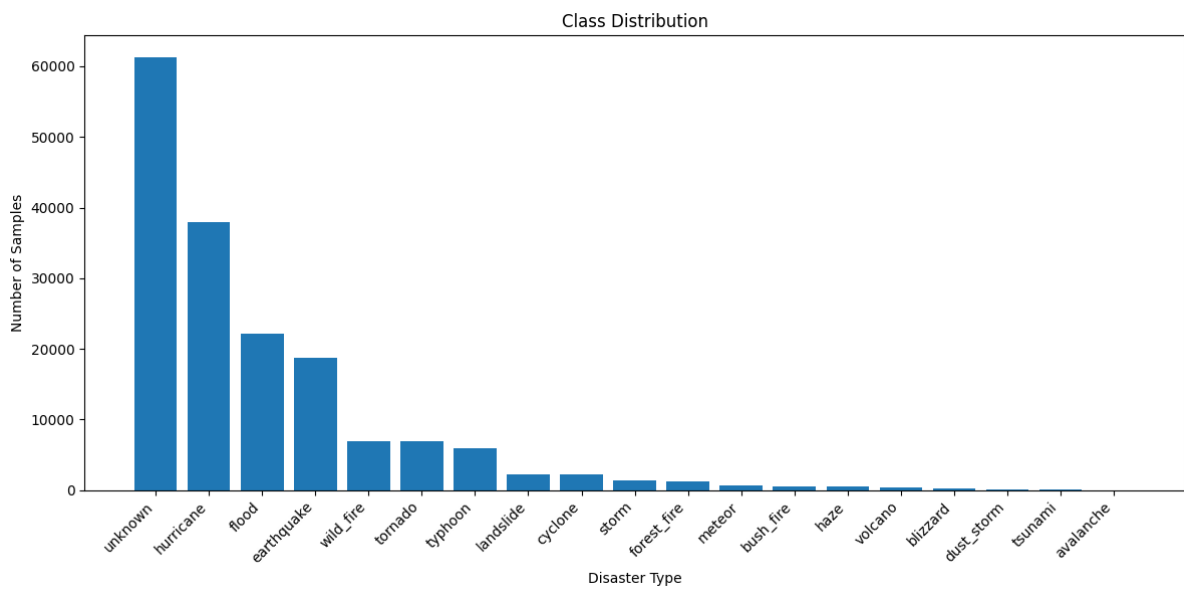
6.3 Text Length Distribution



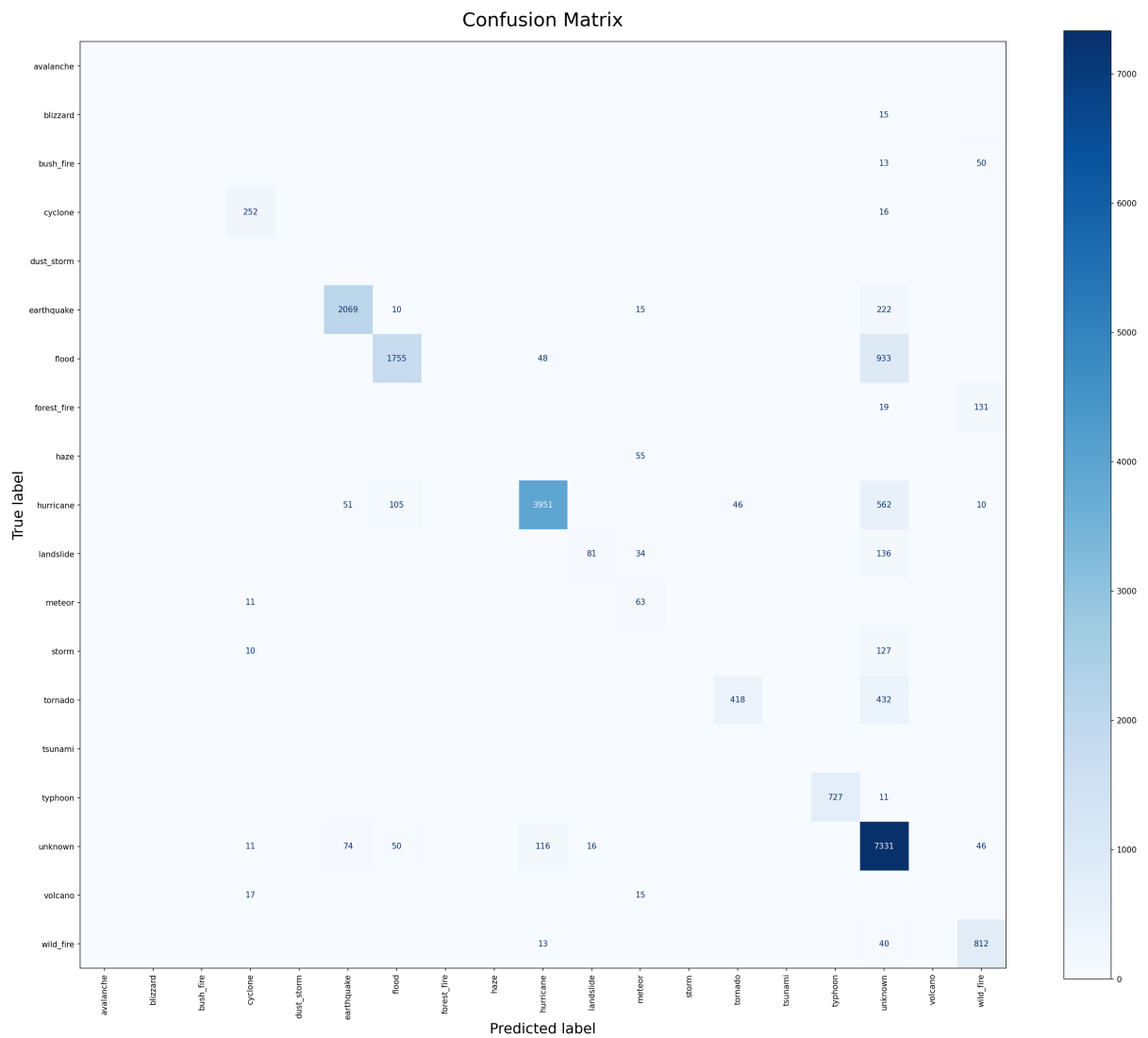
6.4 Training and Validation Loss



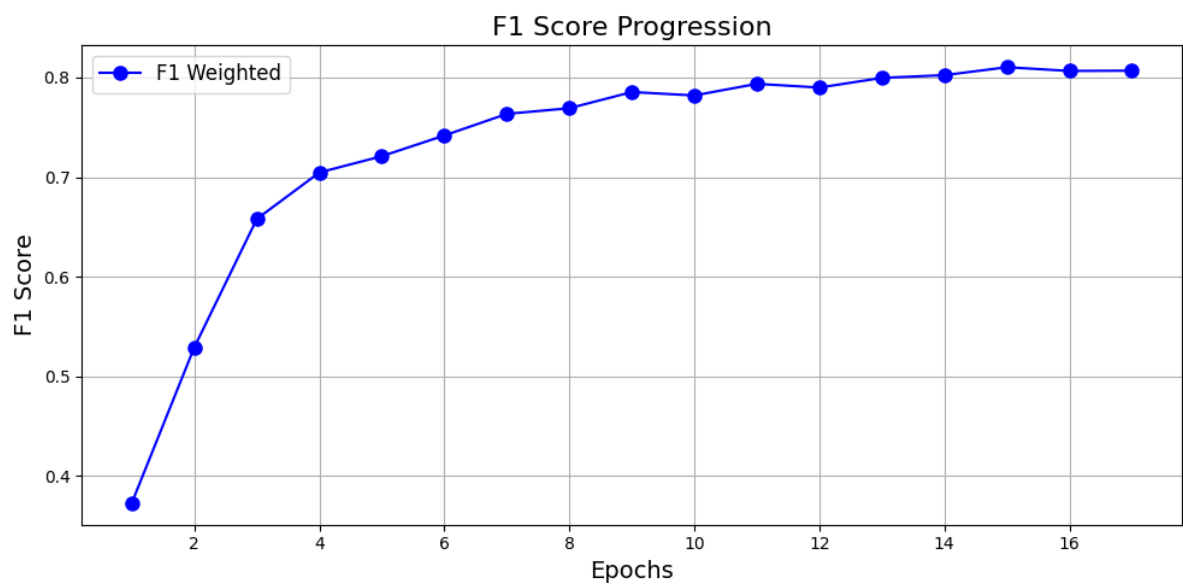
6.5 Class Distribution



6.6 Confusion Matrix

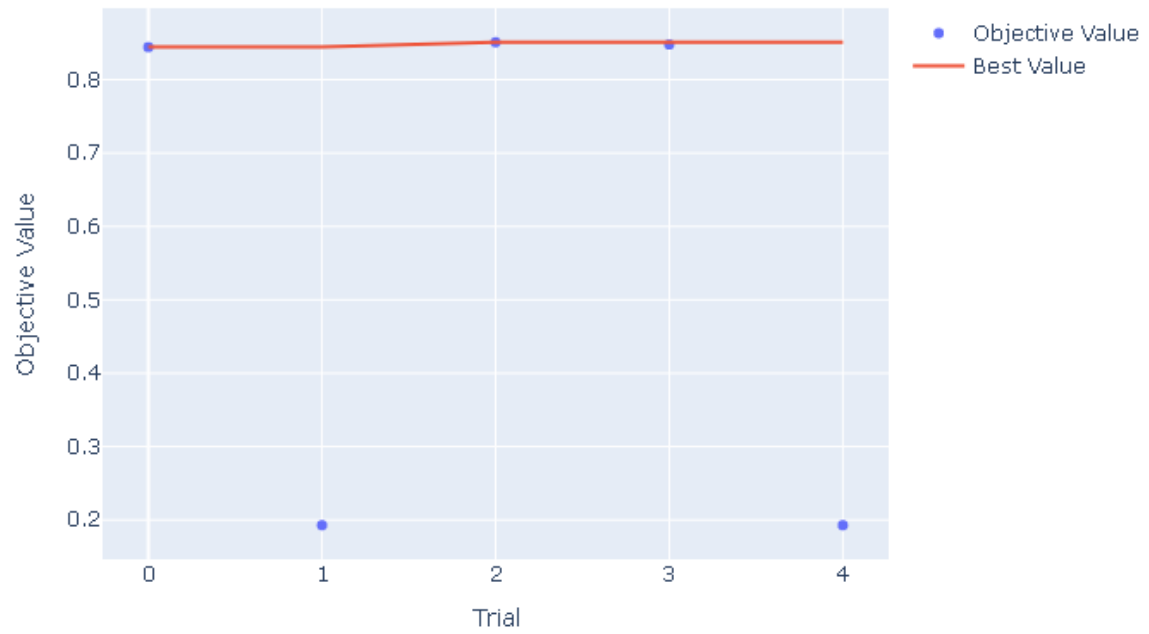


6.7 F1 Score Progression



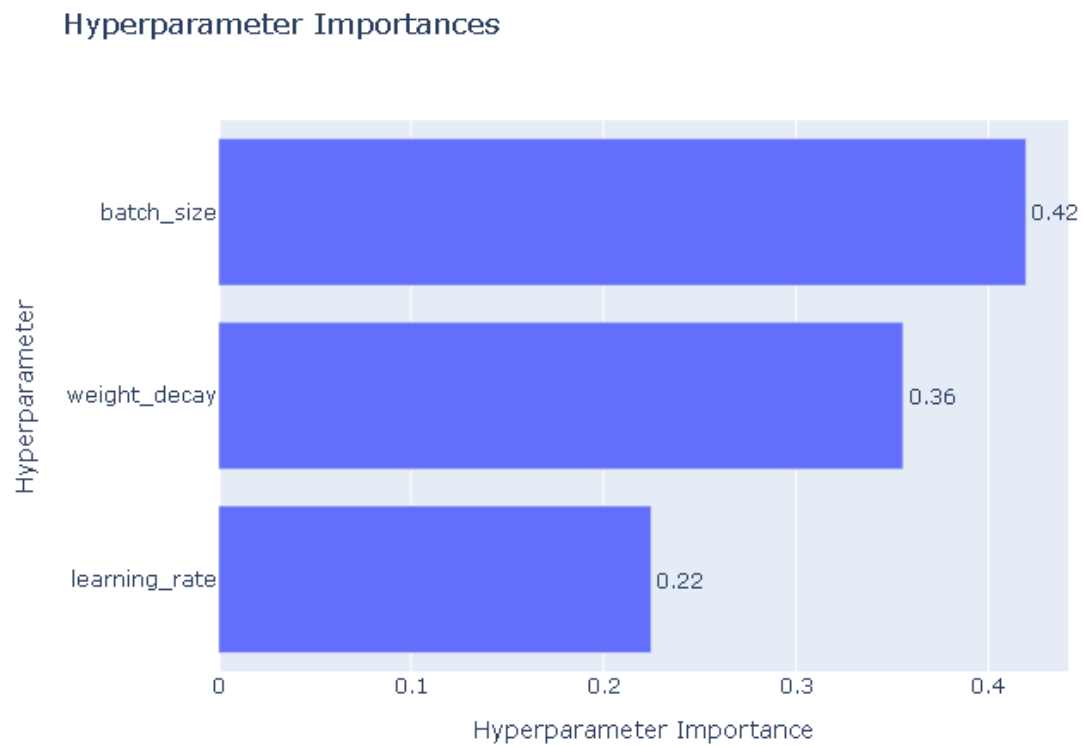
6.8 Optimization History Plot

Optimization History Plot

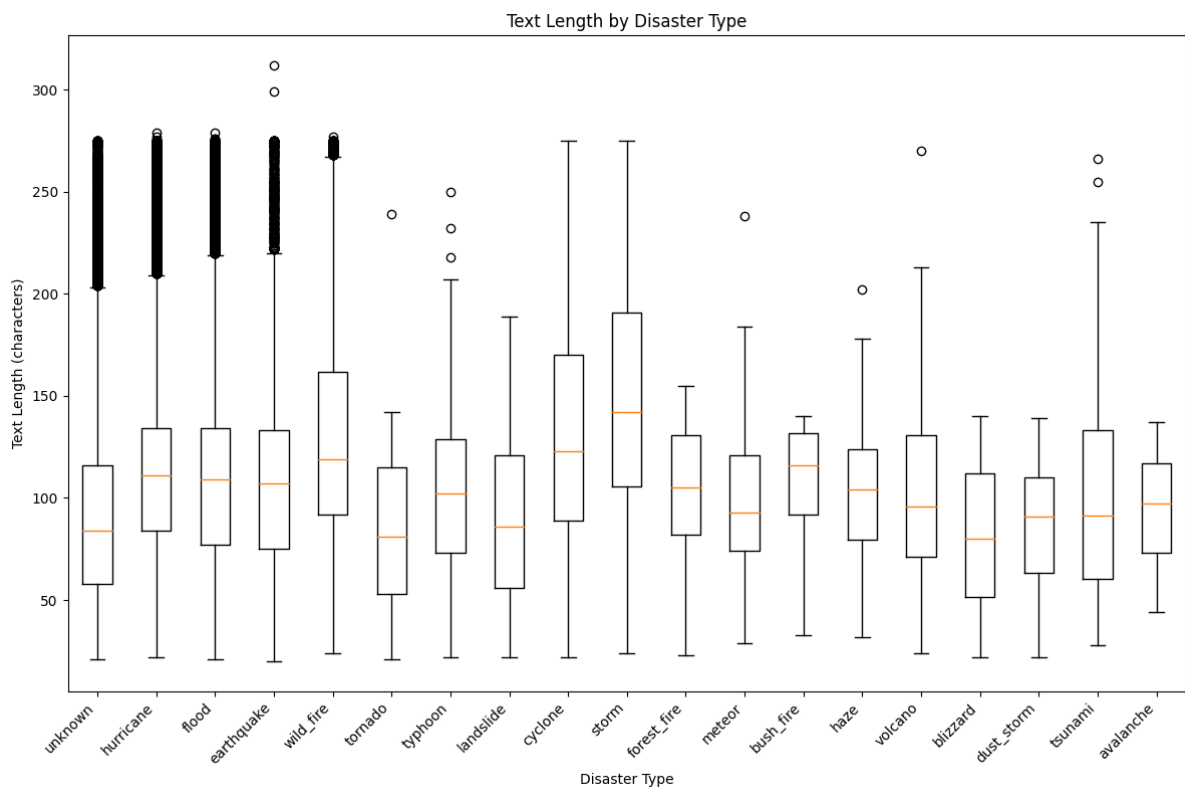


7 Output (XLNET)

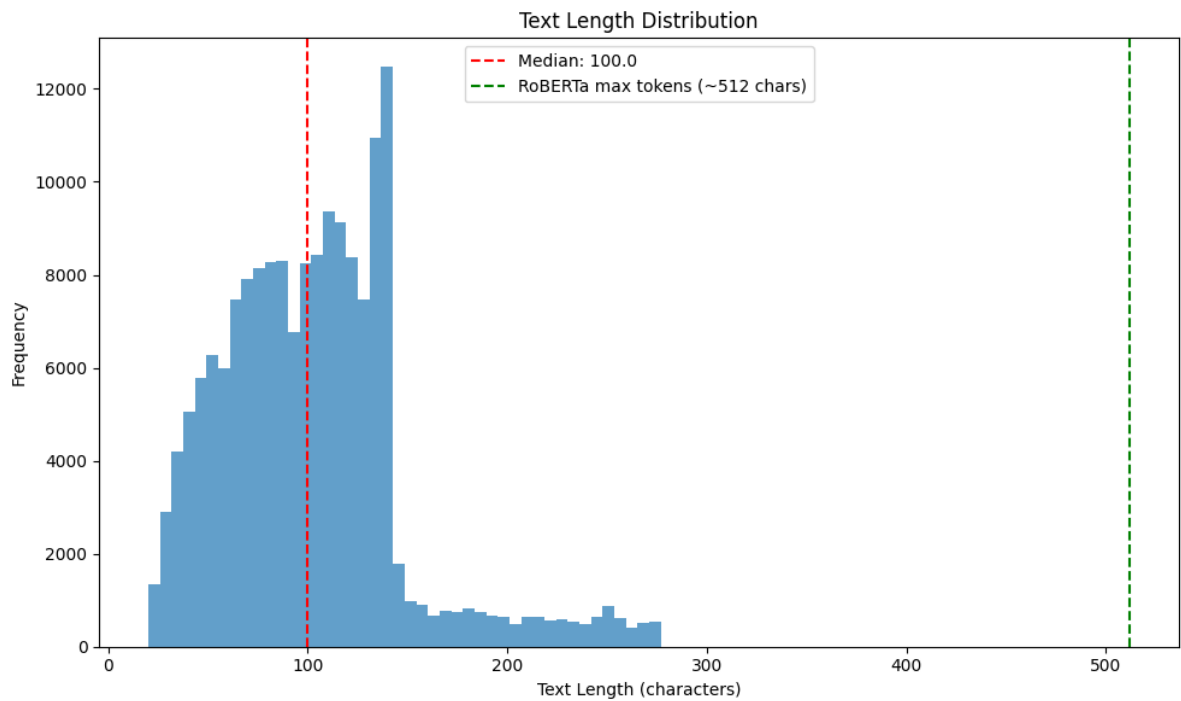
7.1 Hyperparameter Importances



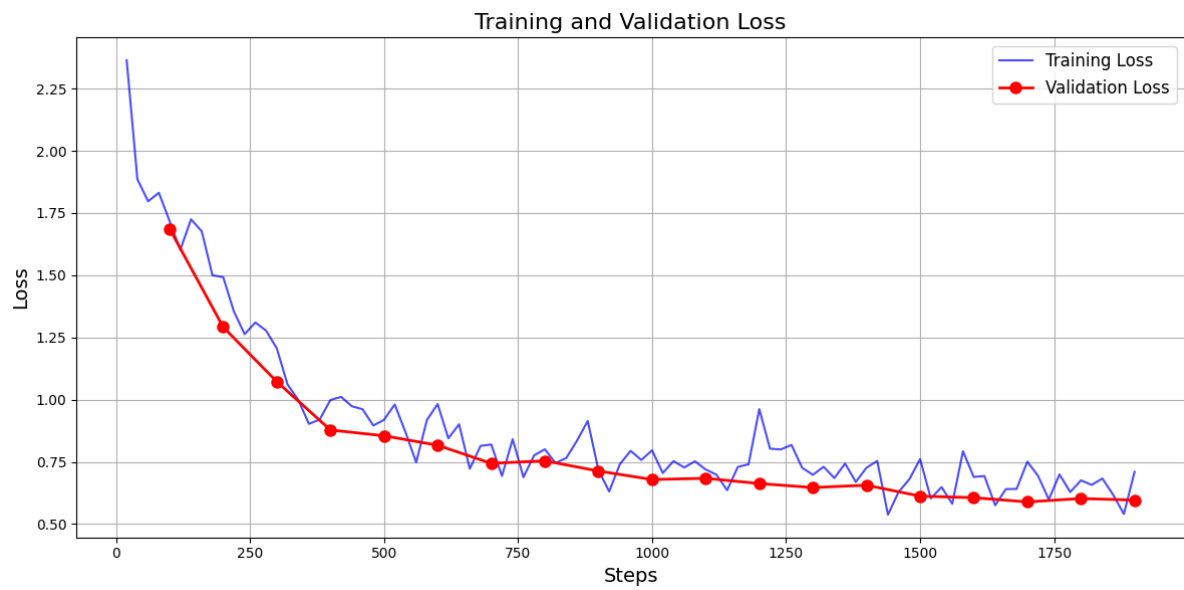
7.2 Text Length by Disaster Type



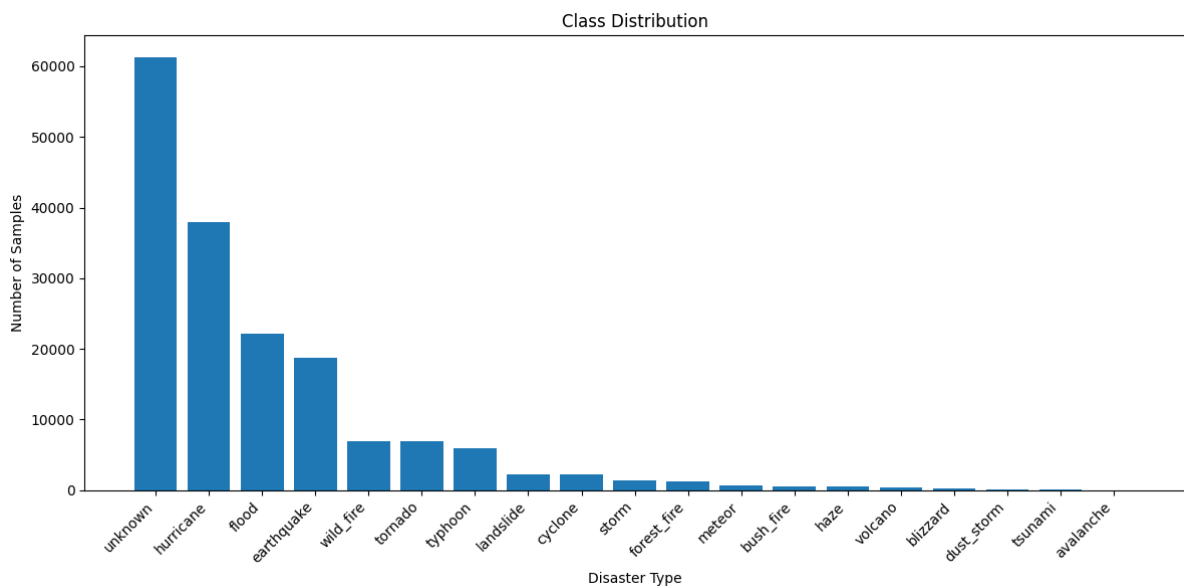
7.3 Text Length Distribution



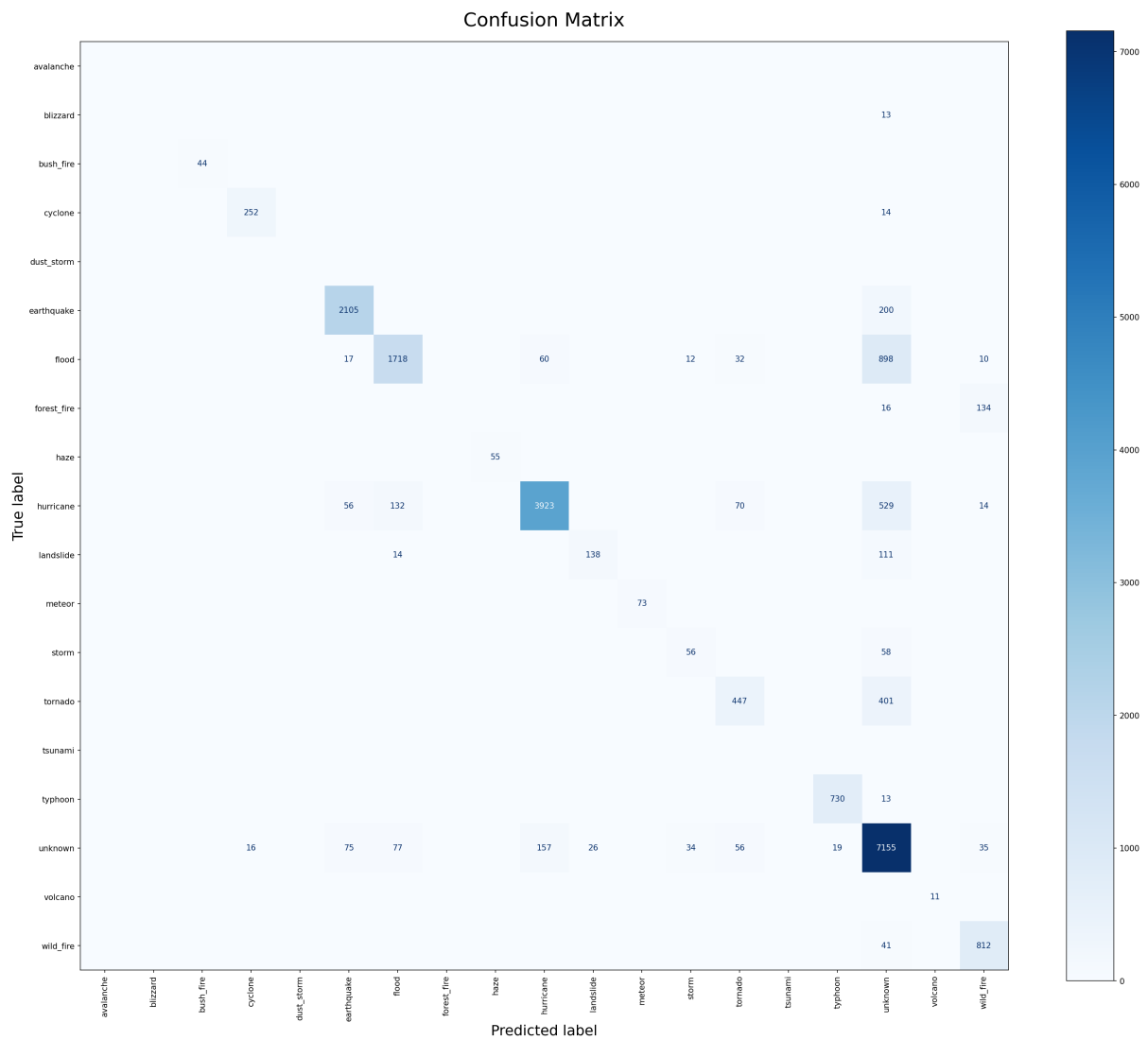
7.4 Training and Validation Loss



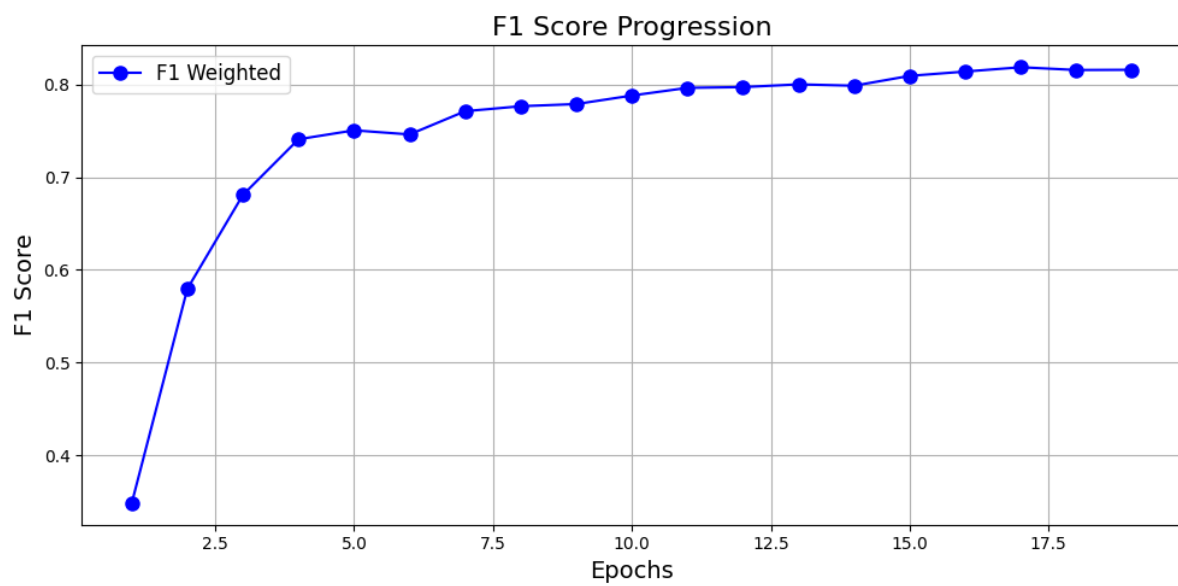
7.5 Class Distribution



7.6 Confusion Matrix



7.7 F1 Score Progression



7.8 Optimization History Plot

Optimization History Plot

