

BERT Finetuning:

Training results

Training Loss	Epoch	Step	Validation Loss	Accuracy	F1	Precision	Recall
0.611	1.0	218	0.4727	0.8170	0.8078	0.8356	0.8170
0.4036	2.0	436	0.4077	0.8557	0.8525	0.8546	0.8557
0.2149	3.0	654	0.4153	0.8711	0.8715	0.8723	0.8711
0.067	4.0	872	0.5430	0.8789	0.8785	0.8784	0.8789

Training hyperparameters

The following hyperparameters were used during training:

- learning_rate: 5e-05
- train_batch_size: 16
- eval_batch_size: 64
- seed: 42
- optimizer: Adam with betas=(0.9,0.999) and epsilon=1e-08
- lr_scheduler_type: linear
- lr_scheduler_warmup_steps: 500
- num_epochs: 4

Fine-tuned version achieves the following results on the evaluation set:

- Loss: 0.5430
- Accuracy: 0.8789
- F1: 0.8785
- Precision: 0.8784
- Recall: 0.8789

Classification report on test dataset:

	precision	recall	f1-score	support
0	0.87	0.82	0.84	110
1	0.89	0.89	0.89	571
2	0.80	0.81	0.80	289
accuracy			0.86	970
macro avg	0.85	0.84	0.85	970
weighted avg	0.86	0.86	0.86	970

RoBERTa Finetuning:

Training results

Training Loss	Epoch	Step	Validation Loss	Accuracy	F1	Precision	Recall
0.3263	1.0	218	0.3427	0.8711	0.8725	0.8797	0.8711
0.377	2.0	436	0.4305	0.8454	0.8458	0.8470	0.8454
0.2243	3.0	654	0.5149	0.8763	0.8756	0.8769	0.8763
0.1564	4.0	872	0.5080	0.8660	0.8658	0.8663	0.8660

Training hyperparameters

The following hyperparameters were used during training:

- learning_rate: 5e-05
- train_batch_size: 16
- eval_batch_size: 64
- seed: 42
- optimizer: Adam with betas=(0.9,0.999) and epsilon=1e-08
- lr_scheduler_type: linear
- lr_scheduler_warmup_steps: 500
- num_epochs: 4

Fine-tuned version achieves the following results on the evaluation set:

- Loss: 0.5080
- Accuracy: 0.8660
- F1: 0.8658
- Precision: 0.8663
- Recall: 0.8660

Classification report on test dataset:

	precision	recall	f1-score	support
0	0.90	0.85	0.87	110
1	0.90	0.91	0.90	571
2	0.84	0.84	0.84	289
accuracy			0.88	970
macro avg	0.88	0.87	0.87	970
weighted avg	0.88	0.88	0.88	970

SVM + CountVectorizer

Classification Report:

	precision	recall	f1-score	support
negative	0.68	0.63	0.65	110
neutral	0.80	0.85	0.82	571
positive	0.68	0.61	0.64	289
accuracy			0.75	970
macro avg	0.72	0.70	0.70	970
weighted avg	0.75	0.75	0.75	970

SVM + TfidfVectorizer

Classification Report:

	precision	recall	f1-score	support
negative	0.85	0.57	0.68	110
neutral	0.76	0.93	0.84	571
positive	0.78	0.54	0.64	289
accuracy			0.77	970
macro avg	0.80	0.68	0.72	970
weighted avg	0.78	0.77	0.76	970