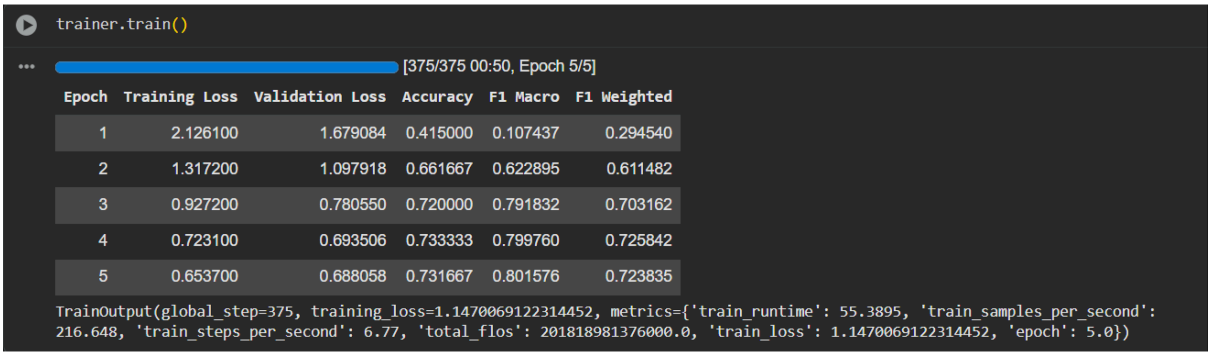
Fine tune distilbert & ALbert

```
MAX_SEQ_LEN = 64
BATCH_SIZE = 32
MODEL_NAME = "distilbert-base-uncased"

# LoRA configuration
lora_config = LoraConfig(
    task_type=TaskType.SEQ_CLS,
    r=4,
    lora_alpha=8,
    lora_dropout=0.1,
    bias="none",
    target_modules=["q_lin", "v_lin"],
)
training_args = TrainingArguments(
    output_dir='./DistilBERT-results',
    num_train_epochs=5,
    per_device_train_batch_size=BATCH_SIZE,
    per_device_eval_batch_size=BATCH_SIZE,
    warmup_steps=200,
    weight_decay=0.01,
    logging_steps=50,
    eval_strategy="epoch",
    save_strategy="epoch",
    load_best_model_at_end=True,
    metric_for_best_model='f1_weighted',
    learning_rate=2e-4,
    save_total_limit=2,
    report_to="none",
)
```

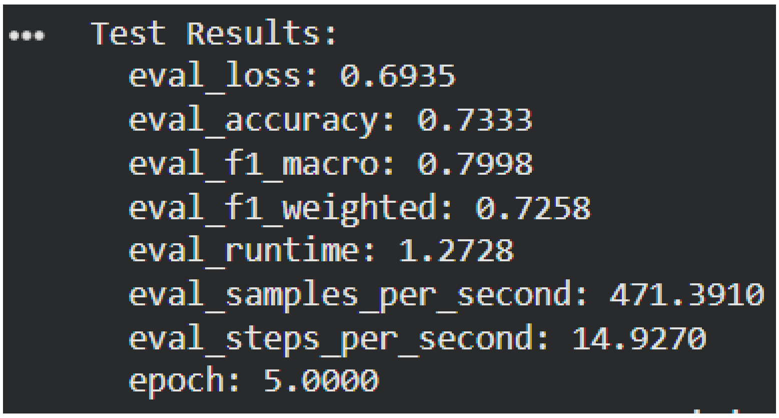


```
trainer.train()
```

[375/375 00:50, Epoch 5/5]

Epoch	Training Loss	Validation Loss	Accuracy	F1 Macro	F1 Weighted
1	2.126100	1.679084	0.415000	0.107437	0.294540
2	1.317200	1.097918	0.661667	0.622895	0.611482
3	0.927200	0.780550	0.720000	0.791832	0.703162
4	0.723100	0.693506	0.733333	0.799760	0.725842
5	0.653700	0.688058	0.731667	0.801576	0.723835

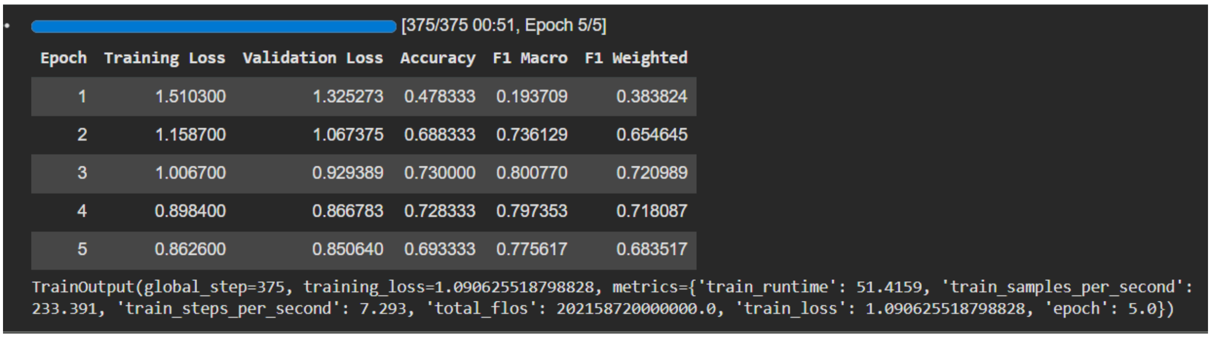
```
TrainOutput(global_step=375, training_loss=1.1470069122314452, metrics={'train_runtime': 55.3895, 'train_samples_per_second':
216.648, 'train_steps_per_second': 6.77, 'total_flos': 201818981376000.0, 'train_loss': 1.1470069122314452, 'epoch': 5.0})
```



```
Test Results:
  eval_loss: 0.6935
  eval_accuracy: 0.7333
  eval_f1_macro: 0.7998
  eval_f1_weighted: 0.7258
  eval_runtime: 1.2728
  eval_samples_per_second: 471.3910
  eval_steps_per_second: 14.9270
  epoch: 5.0000
```

Changes i made

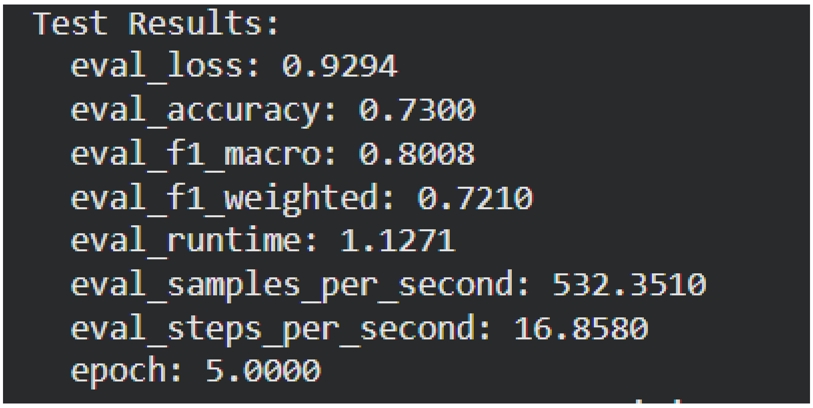
learning_rate: 2e-5
early_stopping_patience=2
r=8
lora alpha = 16
warmup_steps=100



[375/375 00:51, Epoch 5/5]

Epoch	Training Loss	Validation Loss	Accuracy	F1 Macro	F1 Weighted
1	1.510300	1.325273	0.478333	0.193709	0.383824
2	1.158700	1.067375	0.688333	0.736129	0.654645
3	1.006700	0.929389	0.730000	0.800770	0.720989
4	0.898400	0.866783	0.728333	0.797353	0.718087
5	0.862600	0.850640	0.693333	0.775617	0.683517

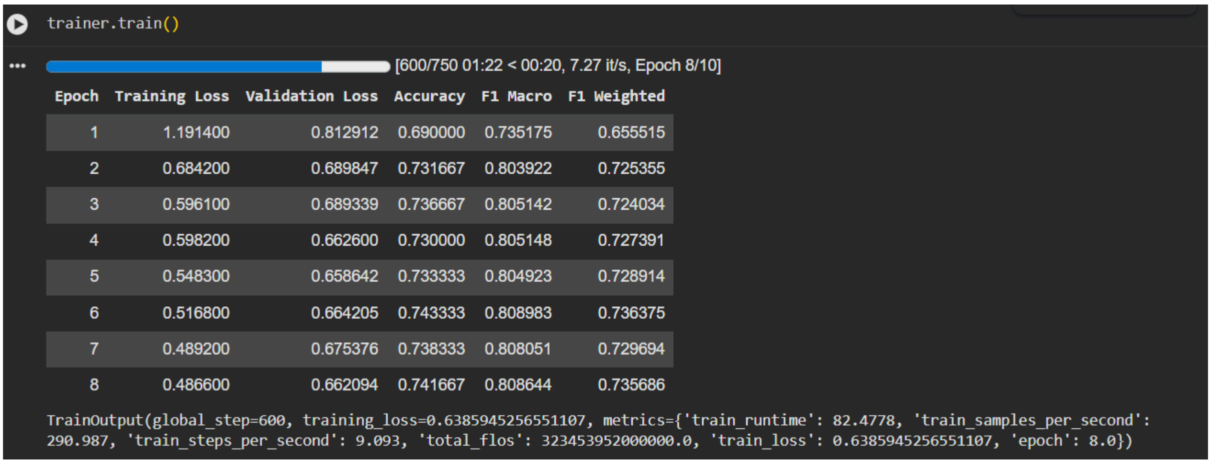
```
TrainOutput(global_step=375, training_loss=1.090625518798828, metrics={'train_runtime': 51.4159, 'train_samples_per_second': 233.391, 'train_steps_per_second': 7.293, 'total_flos': 202158720000000.0, 'train_loss': 1.090625518798828, 'epoch': 5.0})
```



```
Test Results:
  eval_loss: 0.9294
  eval_accuracy: 0.7300
  eval_f1_macro: 0.8008
  eval_f1_weighted: 0.7210
  eval_runtime: 1.1271
  eval_samples_per_second: 532.3510
  eval_steps_per_second: 16.8580
  epoch: 5.0000
```

changes i made

learning_rate: 2e-4
num_train_epochs=10



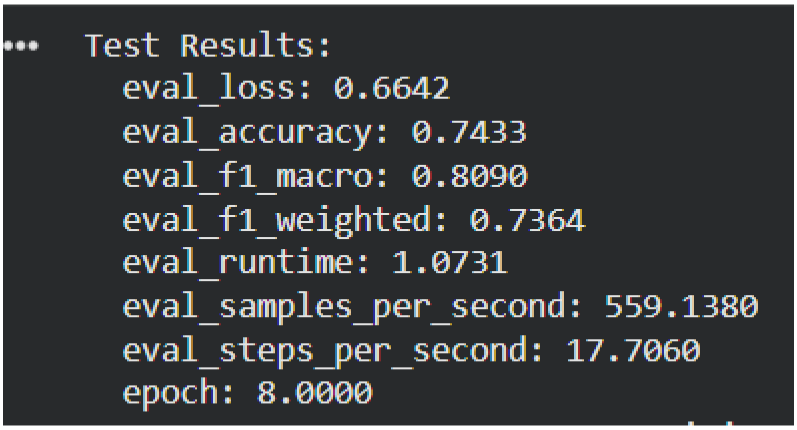
```
trainer.train()
```

[600/750 01:22 < 00:20, 7.27 it/s, Epoch 8/10]

Epoch	Training Loss	Validation Loss	Accuracy	F1 Macro	F1 Weighted
1	1.191400	0.812912	0.690000	0.735175	0.655515
2	0.684200	0.689847	0.731667	0.803922	0.725355
3	0.596100	0.689339	0.736667	0.805142	0.724034
4	0.598200	0.662600	0.730000	0.805148	0.727391
5	0.548300	0.658642	0.733333	0.804923	0.728914
6	0.516800	0.664205	0.743333	0.808983	0.736375
7	0.489200	0.675376	0.738333	0.808051	0.729694
8	0.486600	0.662094	0.741667	0.808644	0.735686

```
TrainOutput(global_step=600, training_loss=0.6385945256551107, metrics={'train_runtime': 82.4778, 'train_samples_per_second':
290.987, 'train_steps_per_second': 9.093, 'total_flos': 323453952000000.0, 'train_loss': 0.6385945256551107, 'epoch': 8.0})
```

Best epoch so far: Epoch 6 with F1 Weighted = **0.7364**



```
Test Results:
  eval_loss: 0.6642
  eval_accuracy: 0.7433
  eval_f1_macro: 0.8090
  eval_f1_weighted: 0.7364
  eval_runtime: 1.0731
  eval_samples_per_second: 559.1380
  eval_steps_per_second: 17.7060
  epoch: 8.0000
```