1012-hw3

Wenxin Zhang

March 2022

1 Reporting results

Report the evaluation metrics and tuned hyperparameters of your best run. Were there any other models that had higher loss but better evaluation accuracy or f1 score? Did the objective value vary a lot across runs?

Solution:

```
== Status ==
Current time: 2022-03-08 03:19:23 (running for 02:13:17.09)
Memory usage on this node: 5.6/29.2 GiB
Using FIFO scheduling algorithm.
Resources requested: 0/8 CPUs, 0/1 GPUs, 0.0/15.15 GiB heap, 0.0/7.57 GiB objects (0.0/1.0 accelerator_type:T4)
Result logdir: /scratch/wz2164/log_dir/_objective_2022-03-08_01-06-06
Number of trials: 5/5 (5 TERMINATED)
+--------------------+------------+-----------------+-----------------+-------------+
| Trial name         | status     | loc             |   learning_rate |   objective |
|--------------------+------------+-----------------+-----------------+-------------|
| _objective_d74c84bf | TERMINATED | 10.144.0.159:5556 |     2.49816e-05 |    0.629856 |
| _objective_d7737f34 | TERMINATED | 10.144.0.159:5557 |     4.80286e-05 |    0.664022 |
| _objective_91959b33 | TERMINATED | 10.144.0.159:5555 |     3.92798e-05 |    0.664037 |
| _objective_498e106a | TERMINATED | 10.144.0.159:5553 |     3.39463e-05 |    0.664029 |
| _objective_03e9b785 | TERMINATED | 10.144.0.159:5558 |     1.23233e-05 |    0.622935 |
+--------------------+------------+-----------------+-----------------+-------------+
```

```
[2m[36m(_objective pid=5558)[0m {'eval_loss': 0.6229350566864014, 'eval_accuracy': 0.7963302752293578, 'eval_f1': 0.8399807784718886,
'eval_precision': 0.8191190253045924, 'eval_recall': 0.8619329388560157, 'eval_runtime': 27.5366, 'eval_samples_per_second': 59.375,
'eval_steps_per_second': 7.445, 'epoch': 3.0}
[2m[36m(_objective pid=5558)[0m {'train_runtime': 1592.2028, 'train_samples_per_second': 17.762, 'train_steps_per_second': 2.221,
'train_loss': 0.4724437016904236, 'epoch': 3.0}
```

Table1: Learning Rate & Eval Loss

Learning Rate	Eval Loss
2.49816e-05	0.629856
4.80286e-05	0.664022
3.92798e-05	0.664037
3.39463e-05	0.664029
1.23233e-05	0.622935

From the table above, we can see that the best learning rate is 1.23233e-05.

Table2: Evaluation of the best run:

Eval Loss	0.6229350566864014
Eval Accuracy	0.7963302752293578
Eval F1	0.8399807784718886
Eval Precision	0.8191190253045924
Eval Recall	0.8619329388560157

Analysis: There aren't any models that had higher loss but better evaluation accuracy or f1 socre. And the objective value did not vary a lot across runs.