

Project 02

1. RNN (LSTM)

Terminal command

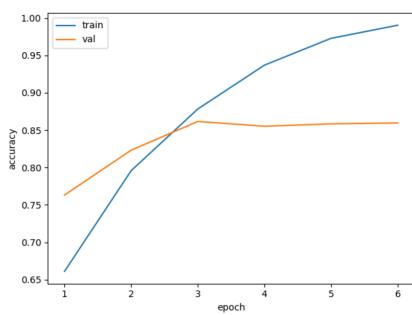
```
python -m src.train --model rnn --run_name rnn_lstm
```

Results

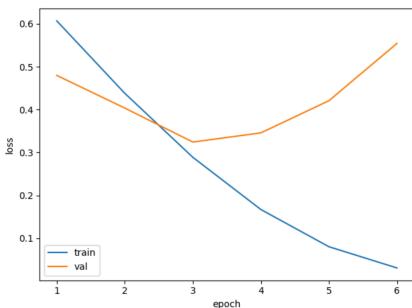
Best

```
{
    "best_val_acc": 0.8616,
    "test_loss": 0.348877844042778,
    "test_acc": 0.85356,
    "best_state_path": "outputs/runs/rnn_lstm_20260131_131945/best_model.pt"
}
```

curves_acc



curves_loss



2. Vanilla transformer

Terminal command

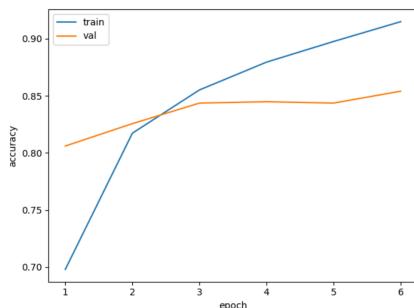
```
python -m src.train --model vanilla --run_name vanilla_tf --lr 3e-4 --dropout 0.1
```

Results

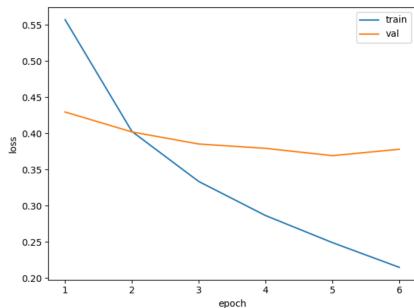
Best

```
{
    "best_val_acc": 0.854,
    "test_loss": 0.40580077671051024,
    "test_acc": 0.83872,
    "best_state_path": "outputs/runs/vanilla_tf_20260131_133546/best_model.pt"
}
```

curves_acc



curves_acc



3. Tweaked transformer

3.1 tweaked_prenorm

Terminal command

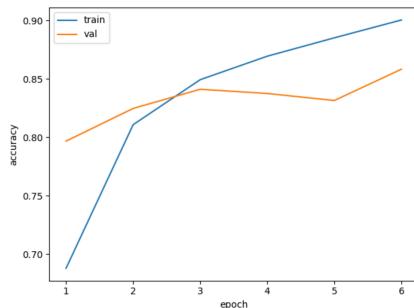
```
python -m src.train --model tweaked_prenorm --run_name tweaked_prenorm --lr 3e-4 --dropout 0.1
```

Results

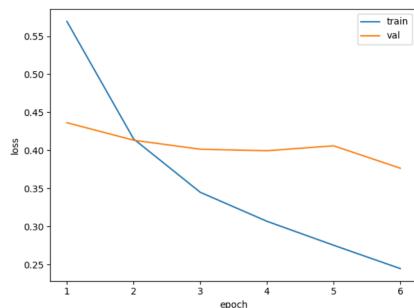
Best

```
{
    "best_val_acc": 0.8584,
    "test_loss": 0.3827622554969788,
    "test_acc": 0.84416,
    "best_state_path": "outputs/runs/tweaked_prenorm_20260131_151138/best_model.pt"
}
```

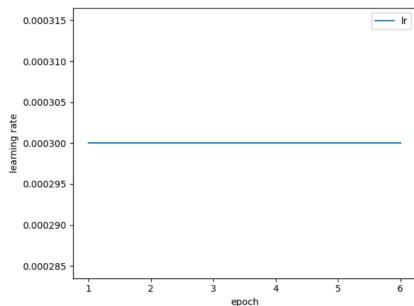
curves_acc



curves_loss



curves_lr



3.2 tweaked_warmup

Terminal command

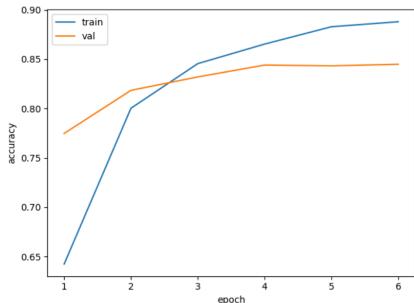
```
python -m src.train --model tweaked_warmup --run_name tweaked_warmup --lr 3e-4 --dropout 0.1 --warmup_ratio 0.1
```

Results

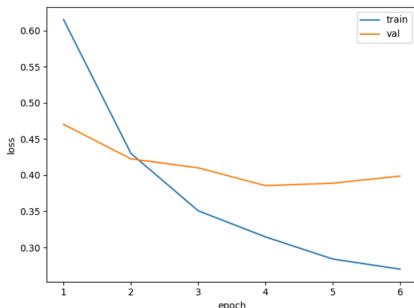
Best

```
{
  "best_val_acc": 0.8448,
  "test_loss": 0.39012223838806154,
  "test_acc": 0.83772,
  "best_state_path": "outputs/runs/tweaked_warmup_20260131_154433/best_model.pt"
}
```

curves_acc



curves_loss



curves_lr

