root@autodl-container-11d445a847-184a60d3:~/hate/hate-speech-detection-survey# python -m src.training.gridsearch

[nltk\_data] Downloading package stopwords to /root/nltk\_data...

[nltk\_data] Package stopwords is already up-to-date!

Device: cuda

hyperparameters keys ['model\_type', 'optimizer\_type', 'loss\_criterion', 'lr', 'epochs', 'batch\_size', 'patience\_es', 'scheduler\_type', 'patience\_lr', 'save\_condition', 'fix\_length', 'context\_size', 'pyramid', 'fcs', 'batch\_norm', 'alpha']

Training to do: 7

prev\_model\_type None

model\_type BasicLSTM

Changing tokenizer...

file loaded and formatted..

data split into train/val/test

field objects created

fields and dataset object created

vocabulary built..

fix\_length: None

batch\_size: 64

dataloaders created..

model\_type: BasicLSTM

optimizer\_type: adam

loss\_criterion: bcelosswithlogits

learning rate: 0.001

epochs: 100

patience\_es: 5

scheduler\_type:

patience\_lr: 3

save\_condition: acc

Model BasicLSTM loaded on cuda

Loss used: BCEWithLogitsLoss()

Optimizer used: Adam (

Parameter Group 0

amsgrad: False

betas: (0.9, 0.999)

eps: 1e-08

lr: 0.001

weight\_decay: 0

)

Scheduler used: None

\*\*TRAINING\*\*

Epoch 0/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 68.83it/s]

train Loss: 0.6076 Acc: 0.4519

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 174.90it/s]

val Loss: 0.5866 Acc: 0.5364

Epoch complete in 2.7s

Epoch 1/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 71.54it/s]

train Loss: 0.4938 Acc: 0.6817

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 224.23it/s]

val Loss: 0.5683 Acc: 0.6440

Epoch complete in 2.5s

Epoch 2/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 72.00it/s]

train Loss: 0.3814 Acc: 0.7867

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 209.64it/s]

val Loss: 0.5987 Acc: 0.6531

Epoch complete in 2.5s

Epoch 3/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 72.88it/s]

train Loss: 0.2588 Acc: 0.8767

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 227.90it/s]

val Loss: 0.6788 Acc: 0.6523

Epoch complete in 2.5s

EarlyStopping counter: 1 out of 5

Epoch 4/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 74.85it/s]

train Loss: 0.1464 Acc: 0.9345

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 225.23it/s]

val Loss: 0.8717 Acc: 0.6793

Epoch complete in 2.4s

Epoch 5/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 73.26it/s]

train Loss: 0.0749 Acc: 0.9714

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 232.81it/s]

val Loss: 1.1314 Acc: 0.6642

Epoch complete in 2.5s

EarlyStopping counter: 1 out of 5

Epoch 6/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 72.40it/s]

train Loss: 0.0607 Acc: 0.9764

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 232.08it/s]

val Loss: 1.2916 Acc: 0.6801

Epoch complete in 2.5s

Epoch 7/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 72.91it/s]

train Loss: 0.0421 Acc: 0.9841

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 225.97it/s]

val Loss: 1.3692 Acc: 0.6693

Epoch complete in 2.5s

EarlyStopping counter: 1 out of 5

Epoch 8/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 73.00it/s]

train Loss: 0.0289 Acc: 0.9890

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 227.01it/s]

val Loss: 1.5501 Acc: 0.6701

Epoch complete in 2.5s

EarlyStopping counter: 2 out of 5

Epoch 9/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 69.82it/s]

train Loss: 0.0309 Acc: 0.9870

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 226.02it/s]

val Loss: 1.4627 Acc: 0.6681

Epoch complete in 2.6s

EarlyStopping counter: 3 out of 5

Epoch 10/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 69.86it/s]

train Loss: 0.0237 Acc: 0.9911

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 227.98it/s]

val Loss: 1.3935 Acc: 0.6773

Epoch complete in 2.6s

EarlyStopping counter: 4 out of 5

Epoch 11/99 | Trainings remaining: 7

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 71.50it/s]

train Loss: 0.0150 Acc: 0.9936

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 226.26it/s]

val Loss: 1.6540 Acc: 0.6729

Epoch complete in 2.5s

EarlyStopping counter: 5 out of 5

Early stopping

Training complete in 0m 30s

Best val Acc: 0.680100

\*\*TESTING\*\*

Loss used: BCEWithLogitsLoss()

Processing batch 860: 100%|██████████████████████████████████████████████████████████████████| 860/860 [00:01<00:00, 562.22it/s]

Test stats - Loss: 1.1637 Acc: 69.31%

Inference on Testset complete in 1.5s

prev\_model\_type BasicLSTM

model\_type BiLSTM

Changing tokenizer...

file loaded and formatted..

data split into train/val/test

field objects created

fields and dataset object created

vocabulary built..

fix\_length: None

batch\_size: 64

dataloaders created..

model\_type: BiLSTM

optimizer\_type: adam

loss\_criterion: bcelosswithlogits

learning rate: 0.001

epochs: 100

patience\_es: 5

scheduler\_type:

patience\_lr: 3

save\_condition: acc

Model BiLSTM loaded on cuda

Loss used: BCEWithLogitsLoss()

Optimizer used: Adam (

Parameter Group 0

amsgrad: False

betas: (0.9, 0.999)

eps: 1e-08

lr: 0.001

weight\_decay: 0

)

Scheduler used: None

\*\*TRAINING\*\*

Epoch 0/99 | Trainings remaining: 6

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:04<00:00, 41.09it/s]

train Loss: 0.6132 Acc: 0.4693

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 157.10it/s]

val Loss: 0.5847 Acc: 0.5454

Epoch complete in 4.3s

Epoch 1/99 | Trainings remaining: 6

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:04<00:00, 39.97it/s]

train Loss: 0.4935 Acc: 0.6839

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 164.42it/s]

val Loss: 0.5456 Acc: 0.6647

Epoch complete in 4.4s

Epoch 2/99 | Trainings remaining: 6

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:04<00:00, 40.10it/s]

train Loss: 0.3716 Acc: 0.8068

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 157.70it/s]

val Loss: 0.5733 Acc: 0.6921

Epoch complete in 4.4s

Epoch 3/99 | Trainings remaining: 6

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:04<00:00, 41.01it/s]

train Loss: 0.2590 Acc: 0.8831

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 135.56it/s]

val Loss: 0.6315 Acc: 0.6731

Epoch complete in 4.4s

EarlyStopping counter: 1 out of 5

Epoch 4/99 | Trainings remaining: 6

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:04<00:00, 40.16it/s]

train Loss: 0.1625 Acc: 0.9348

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 160.84it/s]

val Loss: 0.7916 Acc: 0.6860

Epoch complete in 4.4s

EarlyStopping counter: 2 out of 5

Epoch 5/99 | Trainings remaining: 6

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:04<00:00, 40.17it/s]

train Loss: 0.1022 Acc: 0.9603

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 162.69it/s]

val Loss: 0.9892 Acc: 0.6916

Epoch complete in 4.4s

EarlyStopping counter: 3 out of 5

Epoch 6/99 | Trainings remaining: 6

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:04<00:00, 40.28it/s]

train Loss: 0.0680 Acc: 0.9741

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 165.22it/s]

val Loss: 1.0931 Acc: 0.6866

Epoch complete in 4.4s

EarlyStopping counter: 4 out of 5

Epoch 7/99 | Trainings remaining: 6

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:04<00:00, 40.37it/s]

train Loss: 0.0569 Acc: 0.9789

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 166.63it/s]

val Loss: 1.2275 Acc: 0.6877

Epoch complete in 4.4s

EarlyStopping counter: 5 out of 5

Early stopping

Training complete in 0m 35s

Best val Acc: 0.692100

\*\*TESTING\*\*

Loss used: BCEWithLogitsLoss()

Processing batch 860: 100%|██████████████████████████████████████████████████████████████████| 860/860 [00:01<00:00, 480.85it/s]

Test stats - Loss: 0.5242 Acc: 69.05%

Inference on Testset complete in 1.8s

prev\_model\_type BiLSTM

model\_type HybridCNNLSTM

Changing tokenizer...

file loaded and formatted..

data split into train/val/test

field objects created

fields and dataset object created

vocabulary built..

fix\_length: None

batch\_size: 64

dataloaders created..

model\_type: HybridCNNLSTM

optimizer\_type: adam

loss\_criterion: bcelosswithlogits

learning rate: 0.001

epochs: 100

patience\_es: 5

scheduler\_type:

patience\_lr: 3

save\_condition: acc

Model HybridCNNLSTM loaded on cuda

Loss used: BCEWithLogitsLoss()

Optimizer used: Adam (

Parameter Group 0

amsgrad: False

betas: (0.9, 0.999)

eps: 1e-08

lr: 0.001

weight\_decay: 0

)

Scheduler used: None

\*\*TRAINING\*\*

Epoch 0/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 47.47it/s]

train Loss: 0.6277 Acc: 0.4015

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 175.85it/s]

val Loss: 0.6251 Acc: 0.5397

Epoch complete in 3.8s

Epoch 1/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 51.51it/s]

train Loss: 0.5658 Acc: 0.5431

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 185.84it/s]

val Loss: 0.6558 Acc: 0.5429

Epoch complete in 3.5s

Epoch 2/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 50.40it/s]

train Loss: 0.4591 Acc: 0.7180

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 181.31it/s]

val Loss: 0.5419 Acc: 0.6186

Epoch complete in 3.6s

Epoch 3/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 49.96it/s]

train Loss: 0.3462 Acc: 0.8148

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 183.42it/s]

val Loss: 0.5815 Acc: 0.6705

Epoch complete in 3.6s

Epoch 4/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 49.72it/s]

train Loss: 0.2441 Acc: 0.8838

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 182.97it/s]

val Loss: 0.6928 Acc: 0.6963

Epoch complete in 3.6s

Epoch 5/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 50.99it/s]

train Loss: 0.1641 Acc: 0.9297

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 187.40it/s]

val Loss: 0.8558 Acc: 0.6789

Epoch complete in 3.5s

EarlyStopping counter: 1 out of 5

Epoch 6/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 50.95it/s]

train Loss: 0.1173 Acc: 0.9498

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 186.74it/s]

val Loss: 1.0248 Acc: 0.6804

Epoch complete in 3.5s

EarlyStopping counter: 2 out of 5

Epoch 7/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 50.84it/s]

train Loss: 0.0868 Acc: 0.9653

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 186.86it/s]

val Loss: 1.2417 Acc: 0.6662

Epoch complete in 3.5s

EarlyStopping counter: 3 out of 5

Epoch 8/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 50.53it/s]

train Loss: 0.0549 Acc: 0.9796

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 181.04it/s]

val Loss: 1.3418 Acc: 0.6735

Epoch complete in 3.5s

EarlyStopping counter: 4 out of 5

Epoch 9/99 | Trainings remaining: 5

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:03<00:00, 50.76it/s]

train Loss: 0.0582 Acc: 0.9780

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 180.69it/s]

val Loss: 1.4160 Acc: 0.6705

Epoch complete in 3.5s

EarlyStopping counter: 5 out of 5

Early stopping

Training complete in 0m 36s

Best val Acc: 0.696300

\*\*TESTING\*\*

Loss used: BCEWithLogitsLoss()

Processing batch 860: 100%|██████████████████████████████████████████████████████████████████| 860/860 [00:02<00:00, 400.66it/s]

Test stats - Loss: 0.6425 Acc: 66.45%

Inference on Testset complete in 2.2s

prev\_model\_type HybridCNNLSTM

model\_type PyramidCNN

Changing tokenizer...

file loaded and formatted..

data split into train/val/test

field objects created

fields and dataset object created

vocabulary built..

fix\_length: None

batch\_size: 64

context\_size: 0

pyramid: [256, 256]

fcs: [128]

batch\_norm: 0

alpha: 0.2

dataloaders created..

model\_type: PyramidCNN

optimizer\_type: adam

loss\_criterion: bcelosswithlogits

learning rate: 0.001

epochs: 100

patience\_es: 5

scheduler\_type:

patience\_lr: 3

save\_condition: acc

Context size is 0

Model PyramidCNN loaded on cuda

Loss used: BCEWithLogitsLoss()

Optimizer used: Adam (

Parameter Group 0

amsgrad: False

betas: (0.9, 0.999)

eps: 1e-08

lr: 0.001

weight\_decay: 0

)

Scheduler used: None

\*\*TRAINING\*\*

Epoch 0/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 96.93it/s]

train Loss: 0.6382 Acc: 0.4015

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 283.96it/s]

val Loss: 0.6523 Acc: 0.3956

Epoch complete in 1.9s

Epoch 1/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 100.28it/s]

train Loss: 0.6188 Acc: 0.4041

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 279.39it/s]

val Loss: 0.6154 Acc: 0.3956

Epoch complete in 1.8s

Epoch 2/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 106.81it/s]

train Loss: 0.5317 Acc: 0.5601

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 285.88it/s]

val Loss: 0.5502 Acc: 0.6164

Epoch complete in 1.7s

Epoch 3/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 107.55it/s]

train Loss: 0.4311 Acc: 0.7302

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 287.64it/s]

val Loss: 0.5494 Acc: 0.6026

Epoch complete in 1.7s

EarlyStopping counter: 1 out of 5

Epoch 4/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 107.78it/s]

train Loss: 0.3667 Acc: 0.7869

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 287.60it/s]

val Loss: 0.5410 Acc: 0.6757

Epoch complete in 1.7s

Epoch 5/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 105.80it/s]

train Loss: 0.3056 Acc: 0.8401

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 287.65it/s]

val Loss: 0.5577 Acc: 0.6876

Epoch complete in 1.7s

Epoch 6/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 104.02it/s]

train Loss: 0.2646 Acc: 0.8654

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 278.04it/s]

val Loss: 0.5715 Acc: 0.6877

Epoch complete in 1.8s

Epoch 7/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 103.11it/s]

train Loss: 0.2154 Acc: 0.9002

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 279.01it/s]

val Loss: 0.6302 Acc: 0.6965

Epoch complete in 1.8s

Epoch 8/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 105.23it/s]

train Loss: 0.1894 Acc: 0.9140

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 286.04it/s]

val Loss: 0.7231 Acc: 0.6973

Epoch complete in 1.8s

Epoch 9/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 102.55it/s]

train Loss: 0.1613 Acc: 0.9257

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 282.79it/s]

val Loss: 0.8387 Acc: 0.6926

Epoch complete in 1.8s

EarlyStopping counter: 1 out of 5

Epoch 10/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 104.18it/s]

train Loss: 0.1239 Acc: 0.9486

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 282.35it/s]

val Loss: 0.8731 Acc: 0.6987

Epoch complete in 1.8s

Epoch 11/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 103.60it/s]

train Loss: 0.1136 Acc: 0.9538

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 279.63it/s]

val Loss: 1.0820 Acc: 0.6937

Epoch complete in 1.8s

EarlyStopping counter: 1 out of 5

Epoch 12/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 104.19it/s]

train Loss: 0.1015 Acc: 0.9582

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 284.93it/s]

val Loss: 0.9605 Acc: 0.6761

Epoch complete in 1.8s

EarlyStopping counter: 2 out of 5

Epoch 13/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 104.19it/s]

train Loss: 0.0737 Acc: 0.9714

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 283.59it/s]

val Loss: 1.3993 Acc: 0.6462

Epoch complete in 1.8s

EarlyStopping counter: 3 out of 5

Epoch 14/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 106.40it/s]

train Loss: 0.0858 Acc: 0.9631

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 282.63it/s]

val Loss: 1.2268 Acc: 0.6836

Epoch complete in 1.7s

EarlyStopping counter: 4 out of 5

Epoch 15/99 | Trainings remaining: 4

----------

current lr: 0.001

Processing batch 166: 100%|██████████████████████████████████████████████████████████████████| 166/166 [00:01<00:00, 108.63it/s]

train Loss: 0.0703 Acc: 0.9702

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 285.96it/s]

val Loss: 1.1079 Acc: 0.6745

Epoch complete in 1.7s

EarlyStopping counter: 5 out of 5

Early stopping

Training complete in 0m 28s

Best val Acc: 0.698700

\*\*TESTING\*\*

Loss used: BCEWithLogitsLoss()

Processing batch 860: 100%|██████████████████████████████████████████████████████████████████| 860/860 [00:01<00:00, 559.65it/s]

Test stats - Loss: 0.8398 Acc: 70.00%

Inference on Testset complete in 1.5s

prev\_model\_type PyramidCNN

root@autodl-container-11d445a847-184a60d3:~/hate/hate-speech-detection-survey# python -m src.training.gridsearch --config\_path gridsearch\_config.yml

[nltk\_data] Downloading package stopwords to /root/nltk\_data...

[nltk\_data] Package stopwords is already up-to-date!

Device: cuda

hyperparameters keys ['model\_type', 'optimizer\_type', 'loss\_criterion', 'lr', 'epochs', 'batch\_size', 'patience\_es', 'scheduler\_type', 'patience\_lr', 'save\_condition', 'fix\_length', 'context\_size', 'pyramid', 'fcs', 'batch\_norm', 'alpha']

Training to do: 1

prev\_model\_type None

model\_type HybridLSTMCNN

Changing tokenizer...

file loaded and formatted..

data split into train/val/test

field objects created

fields and dataset object created

vocabulary built..

fix\_length: 20

batch\_size: 64

dataloaders created..

model\_type: HybridLSTMCNN

optimizer\_type: adam

loss\_criterion: bcelosswithlogits

learning rate: 0.001

epochs: 100

patience\_es: 5

scheduler\_type:

patience\_lr: 3

save\_condition: acc

Model HybridLSTMCNN loaded on cuda

Loss used: BCEWithLogitsLoss()

Optimizer used: Adam (

Parameter Group 0

amsgrad: False

betas: (0.9, 0.999)

eps: 1e-08

lr: 0.001

weight\_decay: 0

)

Scheduler used: None

\*\*TRAINING\*\*

Epoch 0/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 65.60it/s]

train Loss: 0.6369 Acc: 0.4015

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 252.16it/s]

val Loss: 0.6448 Acc: 0.3956

Epoch complete in 2.7s

Epoch 1/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 67.53it/s]

train Loss: 0.5918 Acc: 0.4957

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 257.01it/s]

val Loss: 0.5921 Acc: 0.4924

Epoch complete in 2.6s

Epoch 2/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 65.60it/s]

train Loss: 0.5215 Acc: 0.6318

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 254.74it/s]

val Loss: 0.5604 Acc: 0.6345

Epoch complete in 2.7s

Epoch 3/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 65.91it/s]

train Loss: 0.4613 Acc: 0.7099

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 256.79it/s]

val Loss: 0.5643 Acc: 0.6689

Epoch complete in 2.7s

Epoch 4/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 68.31it/s]

train Loss: 0.3878 Acc: 0.7832

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 319.85it/s]

val Loss: 0.6262 Acc: 0.6457

Epoch complete in 2.6s

EarlyStopping counter: 1 out of 5

Epoch 5/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 66.21it/s]

train Loss: 0.2994 Acc: 0.8541

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 254.11it/s]

val Loss: 0.7110 Acc: 0.6800

Epoch complete in 2.7s

Epoch 6/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 65.89it/s]

train Loss: 0.2128 Acc: 0.9065

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 258.24it/s]

val Loss: 0.8163 Acc: 0.6813

Epoch complete in 2.7s

Epoch 7/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 65.87it/s]

train Loss: 0.1497 Acc: 0.9394

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 260.20it/s]

val Loss: 1.0927 Acc: 0.6790

Epoch complete in 2.7s

EarlyStopping counter: 1 out of 5

Epoch 8/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 66.10it/s]

train Loss: 0.1102 Acc: 0.9553

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 259.18it/s]

val Loss: 1.4247 Acc: 0.6786

Epoch complete in 2.7s

EarlyStopping counter: 2 out of 5

Epoch 9/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 66.42it/s]

train Loss: 0.0909 Acc: 0.9644

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 254.72it/s]

val Loss: 1.3738 Acc: 0.6753

Epoch complete in 2.7s

EarlyStopping counter: 3 out of 5

Epoch 10/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 59.15it/s]

train Loss: 0.0741 Acc: 0.9710

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 205.39it/s]

val Loss: 1.6480 Acc: 0.6811

Epoch complete in 3.0s

EarlyStopping counter: 4 out of 5

Epoch 11/99 | Trainings remaining: 1

----------

current lr: 0.001

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:02<00:00, 61.95it/s]

train Loss: 0.0591 Acc: 0.9758

Processing batch 42: 100%|█████████████████████████████████████████████████████████████████████| 42/42 [00:00<00:00, 255.73it/s]

val Loss: 1.7024 Acc: 0.6786

Epoch complete in 2.9s

EarlyStopping counter: 5 out of 5

Early stopping

Training complete in 0m 33s

Best val Acc: 0.681300

\*\*TESTING\*\*

Loss used: BCEWithLogitsLoss()

Processing batch 860: 100%|██████████████████████████████████████████████████████████████████| 860/860 [00:02<00:00, 360.35it/s]

Test stats - Loss: 0.8286 Acc: 67.26%

Inference on Testset complete in 2.4s

Gridsearch complete in 0m 50s

root@autodl-container-11d445a847-184a60d3:~/hate/hate-speech-detection-survey# python -m src.training.gridsearch --config\_path gridsearch\_config.yml

[nltk\_data] Downloading package stopwords to /root/nltk\_data...

[nltk\_data] Package stopwords is already up-to-date!

Device: cuda

hyperparameters keys ['model\_type', 'optimizer\_type', 'loss\_criterion', 'lr', 'epochs', 'batch\_size', 'patience\_es', 'scheduler\_type', 'patience\_lr', 'save\_condition', 'fix\_length', 'context\_size', 'pyramid', 'fcs', 'batch\_norm', 'alpha']

Training to do: 2

prev\_model\_type None

model\_type DistillBertEmotion

Changing tokenizer...

file loaded and formatted..

data split into train/val/test

pad\_index 0

field objects created

fields and dataset object created

vocabulary built..

fix\_length: None

batch\_size: 64

dataloaders created..

model\_type: DistillBertEmotion

optimizer\_type: adam

loss\_criterion: crossentropy

learning rate: 2e-05

epochs: 100

patience\_es: 5

scheduler\_type:

patience\_lr: 3

save\_condition: acc

Model DistillBertEmotion loaded on cuda

Loss used: CrossEntropyLoss()

Optimizer used: Adam (

Parameter Group 0

amsgrad: False

betas: (0.9, 0.999)

eps: 1e-08

lr: 2e-05

weight\_decay: 0

)

Scheduler used: None

\*\*TRAINING\*\*

Epoch 0/99 | Trainings remaining: 2

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.62it/s]

train Loss: 0.4821 Acc: 0.7188

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 35.95it/s]

val Loss: 0.4403 Acc: 0.7720

Epoch complete in 15.5s

Epoch 1/99 | Trainings remaining: 2

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.73it/s]

train Loss: 0.3891 Acc: 0.7996

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 36.11it/s]

val Loss: 0.4403 Acc: 0.7676

Epoch complete in 15.3s

EarlyStopping counter: 1 out of 5

Epoch 2/99 | Trainings remaining: 2

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.45it/s]

train Loss: 0.3254 Acc: 0.8434

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 35.58it/s]

val Loss: 0.4882 Acc: 0.7624

Epoch complete in 15.7s

EarlyStopping counter: 2 out of 5

Epoch 3/99 | Trainings remaining: 2

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.32it/s]

train Loss: 0.2503 Acc: 0.8881

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 36.35it/s]

val Loss: 0.5567 Acc: 0.7638

Epoch complete in 15.8s

EarlyStopping counter: 3 out of 5

Epoch 4/99 | Trainings remaining: 2

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.36it/s]

train Loss: 0.1652 Acc: 0.9351

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 35.34it/s]

val Loss: 0.6394 Acc: 0.7543

Epoch complete in 15.8s

EarlyStopping counter: 4 out of 5

Epoch 5/99 | Trainings remaining: 2

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.74it/s]

train Loss: 0.1117 Acc: 0.9588

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 36.17it/s]

val Loss: 0.6987 Acc: 0.7562

Epoch complete in 15.3s

EarlyStopping counter: 5 out of 5

Early stopping

Training complete in 1m 34s

Best val Acc: 0.772000

\*\*TESTING\*\*

Loss used: CrossEntropyLoss()

Processing batch 860: 100%|██████████████████████████████████████████████████████████████████| 860/860 [00:05<00:00, 161.79it/s]

Test stats - Loss: 0.3881 Acc: 79.35%

Inference on Testset complete in 5.3s

prev\_model\_type DistillBertEmotion

model\_type DistillBert

Changing tokenizer...

file loaded and formatted..

data split into train/val/test

pad\_index 0

field objects created

fields and dataset object created

vocabulary built..

fix\_length: None

batch\_size: 64

dataloaders created..

model\_type: DistillBert

optimizer\_type: adam

loss\_criterion: crossentropy

learning rate: 2e-05

epochs: 100

patience\_es: 5

scheduler\_type:

patience\_lr: 3

save\_condition: acc

Model DistillBert loaded on cuda

Loss used: CrossEntropyLoss()

Optimizer used: Adam (

Parameter Group 0

amsgrad: False

betas: (0.9, 0.999)

eps: 1e-08

lr: 2e-05

weight\_decay: 0

)

Scheduler used: None

\*\*TRAINING\*\*

Epoch 0/99 | Trainings remaining: 1

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.48it/s]

train Loss: 0.4991 Acc: 0.6930

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 35.30it/s]

val Loss: 0.4412 Acc: 0.7614

Epoch complete in 15.7s

Epoch 1/99 | Trainings remaining: 1

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.60it/s]

train Loss: 0.3843 Acc: 0.8009

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 35.92it/s]

val Loss: 0.4385 Acc: 0.7680

Epoch complete in 15.5s

Epoch 2/99 | Trainings remaining: 1

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.40it/s]

train Loss: 0.3125 Acc: 0.8519

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 35.65it/s]

val Loss: 0.4945 Acc: 0.7658

Epoch complete in 15.8s

EarlyStopping counter: 1 out of 5

Epoch 3/99 | Trainings remaining: 1

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.46it/s]

train Loss: 0.2298 Acc: 0.9027

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 36.33it/s]

val Loss: 0.5556 Acc: 0.7586

Epoch complete in 15.7s

EarlyStopping counter: 2 out of 5

Epoch 4/99 | Trainings remaining: 1

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.42it/s]

train Loss: 0.1473 Acc: 0.9442

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 35.23it/s]

val Loss: 0.6408 Acc: 0.7599

Epoch complete in 15.8s

EarlyStopping counter: 3 out of 5

Epoch 5/99 | Trainings remaining: 1

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.68it/s]

train Loss: 0.1069 Acc: 0.9616

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 36.04it/s]

val Loss: 0.7018 Acc: 0.7492

Epoch complete in 15.4s

EarlyStopping counter: 4 out of 5

Epoch 6/99 | Trainings remaining: 1

----------

current lr: 2e-05

Processing batch 166: 100%|███████████████████████████████████████████████████████████████████| 166/166 [00:14<00:00, 11.22it/s]

train Loss: 0.0710 Acc: 0.9758

Processing batch 42: 100%|██████████████████████████████████████████████████████████████████████| 42/42 [00:01<00:00, 35.85it/s]

val Loss: 0.8403 Acc: 0.7508

Epoch complete in 16.0s

EarlyStopping counter: 5 out of 5

Early stopping

Training complete in 1m 51s

Best val Acc: 0.768000

\*\*TESTING\*\*

Loss used: CrossEntropyLoss()

Processing batch 860: 100%|██████████████████████████████████████████████████████████████████| 860/860 [00:04<00:00, 178.79it/s]

Test stats - Loss: 0.3691 Acc: 79.93%

Inference on Testset complete in 4.8s

Gridsearch complete in 4m 38s

root@autodl-container-11d445a847-184a60d3:~/hate/hate-speech-detection-survey#

从你提供的日志信息可以看出，这是一个针对仇恨言论检测模型进行网格搜索（Grid Search）的训练和测试过程。网格搜索是一种超参数调优的方法，通过遍历不同的超参数组合来找到最优的模型配置。下面从多个方面详细分析日志所反映的情况：

### 1. 整体流程概述

整个过程包括以下几个主要步骤：

- \*\*环境准备\*\*：下载并检查NLTK的停用词包，确定使用的计算设备为CUDA（GPU）。

- \*\*超参数设置\*\*：定义了一系列超参数，如模型类型、优化器类型、损失函数、学习率、训练轮数等。

- \*\*模型训练\*\*：对不同的模型类型进行训练，每个模型训练过程中记录训练损失、验证损失和准确率。

- \*\*早停策略\*\*：使用早停策略（Early Stopping）防止过拟合，当验证集上的性能连续多个epoch没有提升时，停止训练。

- \*\*模型测试\*\*：训练完成后，在测试集上评估模型的性能，记录测试损失和准确率。

- \*\*网格搜索完成\*\*：所有模型的训练和测试完成后，输出网格搜索的总耗时。

### 2. 模型训练情况

#### 2.1 不同模型的训练表现

- \*\*BasicLSTM\*\*：

- 训练初期，训练损失较高，准确率较低，随着训练的进行，训练损失迅速下降，准确率大幅提升。

- 验证集上的准确率在训练过程中先上升后趋于稳定，最终在第11个epoch触发早停机制，最佳验证准确率为0.6801。

- 测试集上的准确率为69.31%。

- \*\*BiLSTM\*\*：

- 训练过程与BasicLSTM类似，训练损失逐渐下降，准确率逐渐提高。

- 验证集准确率在第2个epoch达到较高值，随后波动，最终在第7个epoch触发早停，最佳验证准确率为0.6921。

- 测试集准确率为69.05%。

- \*\*HybridCNNLSTM\*\*：

- 训练初期表现不佳，训练损失较高，准确率较低，后续训练损失逐渐降低，准确率有所提升。

- 验证集准确率在第4个epoch达到较高值，之后开始波动，最终在第9个epoch触发早停，最佳验证准确率为0.6963。

- 测试集准确率为66.45%。

- \*\*PyramidCNN\*\*：

- 训练初期训练损失较高，准确率较低，随着训练的进行，损失逐渐下降，准确率逐渐提高。

- 验证集准确率在训练过程中波动上升，最终在第15个epoch触发早停，最佳验证准确率为0.6987。

- 测试集准确率为70.00%。

- \*\*HybridLSTMCNN\*\*：

- 训练初期损失较高，准确率较低，训练过程中损失逐渐下降，准确率逐渐提高。

- 验证集准确率在第6个epoch达到较高值，随后波动，最终在第11个epoch触发早停，最佳验证准确率为0.6813。

- 测试集准确率为67.26%。

- \*\*DistillBertEmotion\*\*：

- 训练过程相对较慢，每个epoch的训练时间较长。训练初期训练损失较高，准确率较低，随着训练的进行，损失逐渐下降，准确率逐渐提高。

- 验证集准确率在第0个epoch达到较高值，随后波动，最终在第5个epoch触发早停，最佳验证准确率为0.7720。

- 测试集准确率为79.35%。

- \*\*DistillBert\*\*：

- 训练过程与DistillBertEmotion类似，每个epoch的训练时间较长。训练损失逐渐下降，准确率逐渐提高。

- 验证集准确率在第1个epoch达到较高值，随后波动，最终在第6个epoch触发早停，最佳验证准确率为0.7680。

- 测试集准确率为79.93%。

#### 2.2 早停策略的作用

早停策略在训练过程中起到了重要作用，避免了模型过拟合。大多数模型在验证集性能不再提升时，早停机制会及时停止训练，节省了计算资源。

### 3. 模型测试结果

从测试集的准确率来看，基于预训练模型的DistillBert和DistillBertEmotion表现明显优于其他传统的深度学习模型（如LSTM、CNN等）。这可能是因为预训练模型在大规模语料上进行了预训练，学习到了丰富的语言知识，能够更好地处理文本数据。

### 4. 计算资源使用情况

- \*\*训练时间\*\*：不同模型的训练时间差异较大，基于预训练模型的DistillBert和DistillBertEmotion每个epoch的训练时间较长，可能是因为模型复杂度较高，需要更多的计算资源。而传统的深度学习模型（如LSTM、CNN等）训练时间相对较短。

- \*\*设备使用\*\*：使用CUDA（GPU）进行计算，大大提高了训练速度。

### 5. 可能的改进方向

- \*\*超参数调整\*\*：虽然进行了网格搜索，但可能还存在更优的超参数组合。可以进一步扩大超参数搜索空间，或者使用更高级的超参数调优方法，如随机搜索、贝叶斯优化等。

- \*\*模型融合\*\*：可以尝试将多个表现较好的模型进行融合，如投票法、堆叠法等，以提高模型的整体性能。

- \*\*数据增强\*\*：可以尝试对训练数据进行增强，如随机替换、插入、删除等操作，以增加数据的多样性，提高模型的泛化能力。

- \*\*模型改进\*\*：可以尝试对现有的模型进行改进，如调整模型结构、增加正则化项等，以提高模型的性能。