



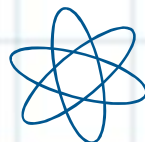
CO₂

FR-Miner论文复现

专业：软件工程

学生/庄毅非

导师/万志远





第一节：运行细节

运行平台： windows docker desktop

运行过程：

1. 安装docker desktop for windows
2. 在powershell中运行docker pull zzyo/frminer:v1
拉取所需镜像
3. 运行docker run -it zzyo/frminer:v1 /bin/bash
创建对应容器，进入home目录。
4. 输入allennlp train FRMiner/config.json -s FRMiner/out/ -f
--include-package FRMiner运行训练程序

程序运行中的一个截图

```
root@86cbb24473b:/home
2022-09-30 04:37:54.593 - INFO - allennlp.training.trainer - Validating
accuracy: 0.7480, precision: 0.7279, recall: 0.7857, fscore: 0.7557, s_precision: 0.8303, s_recall: 0.8709, s_fmeasure: 0.8548, loss: 0.5665 || 100% ##### 8/8 [00:04:00.00, 1.961it/s]
2022-09-30 04:37:58.687 - INFO - allennlp.training.tensorboard.writer - Training Validation
2022-09-30 04:37:58.688 - INFO - allennlp.training.tensorboard.writer - accuracy 0.883 0.748
2022-09-30 04:37:58.688 - INFO - allennlp.training.tensorboard.writer - fscore 0.884 0.756
2022-09-30 04:37:58.689 - INFO - allennlp.training.tensorboard.writer - loss 0.358 0.566
2022-09-30 04:37:58.690 - INFO - allennlp.training.tensorboard.writer - s_precision 0.839 0.839
2022-09-30 04:37:58.690 - INFO - allennlp.training.tensorboard.writer - s_recall 25293.064 N/A
2022-09-30 04:37:58.691 - INFO - allennlp.training.tensorboard.writer - s_fmeasure 0.855 0.871
2022-09-30 04:37:58.692 - INFO - allennlp.training.tensorboard.writer - recall 0.891 0.786
2022-09-30 04:37:58.692 - INFO - allennlp.training.tensorboard.writer - precision 0.877 0.728
2022-09-30 04:37:58.718 - INFO - allennlp.training.checkpointer - Best validation performance so far. Copying weights to 'FRMiner/out/best.tl'.
2022-09-30 04:37:58.723 - INFO - allennlp.training.trainer - Epoch duration: 0:07:23.492454
2022-09-30 04:37:58.723 - INFO - allennlp.training.trainer - Estimated training time remaining: 0:13:30
2022-09-30 04:37:58.723 - INFO - allennlp.training.trainer - Epoch 8/9
2022-09-30 04:37:58.725 - INFO - allennlp.training.trainer - Peak CPU memory usage MB: 25293.064
2022-09-30 04:37:59.277 - INFO - allennlp.training.trainer - Training
accuracy: 0.8818, precision: 0.8782, recall: 0.8903, fscore: 0.8842, s_precision: 0.8402, s_recall: 0.8715, s_fmeasure: 0.8535, loss: 0.3598 || 100% ##### 99/475 [01:24:03:58, 1.581it/s] 2022-09-30 04:39:24.791 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8828, precision: 0.8785, recall: 0.8903, fscore: 0.8840, s_precision: 0.8403, s_recall: 0.8724, s_fmeasure: 0.8536, loss: 0.3599 || 43 ##### 199/475 [02:46:04:12, 1.601it/s] 2022-09-30 04:40:46.589 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8833, precision: 0.8715, recall: 0.8903, fscore: 0.8840, s_precision: 0.8415, s_recall: 0.8734, s_fmeasure: 0.8570, loss: 0.3592 || 43 ##### 299/475 [04:09:02:07, 1.581it/s] 2022-09-30 04:42:38.970 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8842, precision: 0.8764, recall: 0.8941, fscore: 0.8852, s_precision: 0.8425, s_recall: 0.8732, s_fmeasure: 0.8576, loss: 0.3582 || 84 ##### 399/475 [05:39:02:20, 1.06it/s] 2022-09-30 04:43:40.095 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8847, precision: 0.8768, recall: 0.8951, fscore: 0.8858, s_precision: 0.8432, s_recall: 0.8738, s_fmeasure: 0.8582, loss: 0.3577 || 100% ##### 475/475 [06:45:00:00, 1.171it/s]
2022-09-30 04:44:44.304 - INFO - allennlp.training.trainer - Validating
accuracy: 0.7421, precision: 0.7228, recall: 0.7857, fscore: 0.7529, s_precision: 0.8430, s_recall: 0.8736, s_fmeasure: 0.8580, loss: 0.5665 || 100% ##### 8/8 [00:02:00.00, 2.911it/s]
2022-09-30 04:44:47.082 - INFO - allennlp.training.tensorboard.writer - Training Validation
2022-09-30 04:44:47.085 - INFO - allennlp.training.tensorboard.writer - accuracy 0.883 0.742
2022-09-30 04:44:47.086 - INFO - allennlp.training.tensorboard.writer - fscore 0.886 0.753
2022-09-30 04:44:47.086 - INFO - allennlp.training.tensorboard.writer - loss 0.358 0.567
2022-09-30 04:44:47.087 - INFO - allennlp.training.tensorboard.writer - s_precision 0.843 0.843
2022-09-30 04:44:47.087 - INFO - allennlp.training.tensorboard.writer - s_recall 25293.064 N/A
2022-09-30 04:44:47.088 - INFO - allennlp.training.tensorboard.writer - s_fmeasure 0.854 0.874
2022-09-30 04:44:47.088 - INFO - allennlp.training.tensorboard.writer - recall 0.893 0.858
2022-09-30 04:44:47.089 - INFO - allennlp.training.tensorboard.writer - precision 0.895 0.786
2022-09-30 04:44:47.089 - INFO - allennlp.training.tensorboard.writer - precision 0.877 0.723
2022-09-30 04:44:47.094 - INFO - allennlp.training.checkpointer - Best validation performance so far. Copying weights to 'FRMiner/out/best.tl'.
2022-09-30 04:44:47.091 - INFO - allennlp.training.trainer - Epoch duration: 0:06:48.304498
2022-09-30 04:44:47.092 - INFO - allennlp.training.trainer - Estimated training time remaining: 0:06:45
2022-09-30 04:44:47.092 - INFO - allennlp.training.trainer - Epoch 9/9
2022-09-30 04:44:47.094 - INFO - allennlp.training.trainer - Peak CPU memory usage MB: 25293.064
2022-09-30 04:44:47.688 - INFO - allennlp.training.trainer - Training
accuracy: 0.8819, precision: 0.8880, recall: 0.8995, fscore: 0.8937, s_precision: 0.8432, s_recall: 0.8738, s_fmeasure: 0.8582, loss: 0.3516 || 36/5 24/475 [00:29:08:44, 1.16it/s] 2022-09-30 04:45:17.518 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8824, precision: 0.8770, recall: 0.8920, fscore: 0.8896, s_precision: 0.8440, s_recall: 0.8743, s_fmeasure: 0.8583, loss: 0.3520 || 28 ##### 124/475 [01:50:04:12, 1.221it/s] 2022-09-30 04:46:46.598 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8829, precision: 0.8779, recall: 0.8926, fscore: 0.8900, s_precision: 0.8447, s_recall: 0.8754, s_fmeasure: 0.8594, loss: 0.3523 || 47 ##### 224/475 [03:35:03:51, 1.081it/s] 2022-09-30 04:48:24.650 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8832, precision: 0.8745, recall: 0.8923, fscore: 0.8906, s_precision: 0.8453, s_recall: 0.8749, s_fmeasure: 0.8598, loss: 0.3560 || 68 ##### 324/475 [05:05:02:26, 1.031it/s] 2022-09-30 04:49:24.068 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8822, precision: 0.8740, recall: 0.8926, fscore: 0.8920, s_precision: 0.8457, s_recall: 0.8754, s_fmeasure: 0.8603, loss: 0.3573 || 89 ##### 424/475 [06:38:00:48, 1.061it/s] 2022-09-30 04:51:27.114 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8813, precision: 0.8735, recall: 0.8917, fscore: 0.8825, s_precision: 0.8459, s_recall: 0.8754, s_fmeasure: 0.8604, loss: 0.3580 || 100% ##### 475/475 [07:24:00:00, 1.071it/s]
2022-09-30 04:52:12.617 - INFO - allennlp.training.trainer - Validating
accuracy: 0.7480, precision: 0.7228, recall: 0.7857, fscore: 0.7557, s_precision: 0.8457, s_recall: 0.8753, s_fmeasure: 0.8602, loss: 0.5674 || 100% ##### 8/8 [00:04:00.00, 1.781it/s]
2022-09-30 04:52:17.127 - INFO - allennlp.training.tensorboard.writer - Training Validation
2022-09-30 04:52:17.133 - INFO - allennlp.training.tensorboard.writer - accuracy 0.881 0.746
2022-09-30 04:52:17.136 - INFO - allennlp.training.tensorboard.writer - fscore 0.883 0.756
2022-09-30 04:52:17.136 - INFO - allennlp.training.tensorboard.writer - loss 0.358 0.567
2022-09-30 04:52:17.138 - INFO - allennlp.training.tensorboard.writer - s_precision 0.846 0.846
2022-09-30 04:52:17.137 - INFO - allennlp.training.tensorboard.writer - s_recall 25293.064 N/A
2022-09-30 04:52:17.137 - INFO - allennlp.training.tensorboard.writer - s_fmeasure 0.857 0.872
2022-09-30 04:52:17.138 - INFO - allennlp.training.tensorboard.writer - recall 0.890 0.860
2022-09-30 04:52:17.139 - INFO - allennlp.training.tensorboard.writer - precision 0.892 0.786
2022-09-30 04:52:17.139 - INFO - allennlp.training.tensorboard.writer - precision 0.874 0.728
2022-09-30 04:52:17.186 - INFO - allennlp.training.checkpointer - Best validation performance so far. Copying weights to 'FRMiner/out/best.tl'.
2022-09-30 04:52:17.188 - INFO - allennlp.training.trainer - Epoch duration: 0:07:30.365652
2022-09-30 04:52:17.189 - INFO - allennlp.training.checkpointer - Loading best weights
2022-09-30 04:52:17.259 - INFO - allennlp.models.archival - archiving weights and vocabulary to FRMiner/out/model.tar.gz
2022-09-30 04:52:17.480 - INFO - allennlp.common.util - Metrics: {
  "best_epoch": 9,
  "peak_cpu_memory_MB": 25293.064,
  "training_duration": "1:08:22.380769",
  "training_start_epoch": 0,
  "training_epochs": 9,
  "epoch": 9,
  "training_accuracy": 0.881291172595204,
  "training_precision": 0.876157251358602,
}
```



第二节：困难和解决方法

1. 计算机配置不足

因为我的电脑是intel mac，无法直接运行对应的训练程序，所以只能够在docker中运行该训练程序，但是由于我电脑的内存不足，训练程序总是在第一个epoch的训练进度接近10%的时候被docker后台内核kill掉，所以我的电脑无法运行训练程序。经过向机房的老师申请，我申请到了一台有32g内存的win10专业版电脑主机，所以lab2的训练过程都是在该主机上完成的。

2. 网络连接不佳

在拉取内核镜像的时候，因为使用的是机房中的主机，而机房中的主机对外网的访问能力较弱，所以镜像下载的网速非常慢（只有几十kb每秒），经过计算，需要两天的时间才能够将运行镜像pull下来，这是不可接受的。所以我在自己的电脑上配置好vpn之后，将镜像pull到下来，随后使用docker save命令制作对应的镜像tar包，使用u盘将该tar包移动到主机上之后使用docker load命令直接加载制作好的docker镜像，避免了联网下载所造成的速度过慢的问题。



第二节：困难和解决方法

3. 训练镜像本身问题

在运行 `allennlp train FRMiner/config.json -s FRMiner/out/ -f --include-package FRMiner` 命令的时候，总是会在初始化的时候失败，提示找不到openssl库。查询了stackoverflow之后得知解决方法：从openssl官方仓库下载对应版本的源代码，在镜像中进行编译安装，同时配置好环境变量PATH。（这里不能使用apt直接安装openssl，因为新版本的openssl似乎和镜像中的allennlp不太兼容，使用apt安装的openssl会直接报错，提示函数定义不存在。

4. 所需内存过大

在使用机房的32g电脑进行第一次训练的时候，训练仍然在90%多的时候被kill掉了，任务管理器中显示系统总使用的内存最多的时候达到了31.9G，导致了docker容器内存不足，训练中止。于是我将windows defender，chrome后台更新程序等和训练无关的进程关闭之后，之后训练消耗的最大内存仅仅只有30.9G，勉强达到了要求，成功进行了训练。



第三节：训练过程和结果展示

训练过程截图

```
root@88ecbb24473b:/home#
2022-09-30 04:23:47.588 INFO - allennlp.training.tensorboard_writer - accuracy 0.880 0.718
2022-09-30 04:23:47.590 INFO - allennlp.training.tensorboard_writer - f_score 0.882 0.732
2022-09-30 04:23:47.591 INFO - allennlp.training.tensorboard_writer - loss 0.383 0.885
2022-09-30 04:23:47.591 INFO - allennlp.training.tensorboard_writer - s_precision 0.880 0.829
2022-09-30 04:23:47.592 INFO - allennlp.training.tensorboard_writer - cpu_memory_MB 25293.064 N/A
2022-09-30 04:23:47.592 INFO - allennlp.training.tensorboard_writer - s_recall 0.885 0.865
2022-09-30 04:23:47.593 INFO - allennlp.training.tensorboard_writer - s_fmeasure 0.847 0.846
2022-09-30 04:23:47.593 INFO - allennlp.training.tensorboard_writer - recall 0.892 0.770
2022-09-30 04:23:47.594 INFO - allennlp.training.tensorboard_writer - precision 0.871 0.698
2022-09-30 04:23:47.616 INFO - allennlp.training.checkpointer - Best validation performance so far. Copying weights to 'FBIuser/out/best.tl'.
2022-09-30 04:23:47.621 INFO - allennlp.training.trainer - Epoch duration: 0:07:30.307379
2022-09-30 04:23:47.622 INFO - allennlp.training.trainer - Estimated training time remaining: 0:26:35
2022-09-30 04:23:47.622 INFO - allennlp.training.trainer - Epoch 6/9
2022-09-30 04:23:47.623 INFO - allennlp.training.trainer - Peak CPU memory usage MB: 25293.064
2022-09-30 04:23:48.065 INFO - allennlp.training.trainer - Training
accuracy: 0.8941, precision: 0.8825, recall: 0.9039, f_score: 0.8984, s_precision: 0.8302, s_recall: 0.8652, s_fmeasure: 0.8473, loss: 0.3511 | 100% ##### 49/475 [04:41:05:18, 1.341t/s] 2022-09-30 04:24:30.739 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8936, precision: 0.8803, recall: 0.8907, f_score: 0.8864, s_precision: 0.8314, s_recall: 0.8664, s_fmeasure: 0.8485, loss: 0.3547 | 31% ### 149/475 [04:22:03:48, 1.481t/s] 2022-09-30 04:25:51.577 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8882, precision: 0.8757, recall: 0.8907, f_score: 0.8831, s_precision: 0.8324, s_recall: 0.8668, s_fmeasure: 0.8492, loss: 0.3593 | 52% ##### 249/475 [04:22:02:31, 1.491t/s] 2022-09-30 04:27:11.408 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8796, precision: 0.8757, recall: 0.8866, f_score: 0.8811, s_precision: 0.8337, s_recall: 0.8673, s_fmeasure: 0.8501, loss: 0.3605 | 73% ##### 349/475 [04:45:02:11, 1.046t/s] 2022-09-30 04:28:33.557 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8806, precision: 0.8751, recall: 0.8823, f_score: 0.8812, s_precision: 0.8343, s_recall: 0.8681, s_fmeasure: 0.8511, loss: 0.3604 | 85% ##### 449/475 [04:29:06:38, 1.429t/s] 2022-09-30 04:30:12.183 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8734, precision: 0.8748, recall: 0.8855, f_score: 0.8801, s_precision: 0.8356, s_recall: 0.8681, s_fmeasure: 0.8512, loss: 0.3606 | 100% ##### 475/475 [04:44:00:00, 1.171t/s]
2022-09-30 04:30:32.669 INFO - allennlp.training.trainer - Validating
accuracy: 0.7282, precision: 0.7480, recall: 0.7608, f_score: 0.7378, s_precision: 0.8347, s_recall: 0.8678, s_fmeasure: 0.8509, loss: 0.5809 | 100% ##### 8/8 [00:02:00:00, 1.151t/s]
2022-09-30 04:30:35.212 INFO - allennlp.training.tensorboard_writer - Training Validation
2022-09-30 04:30:35.213 INFO - allennlp.training.tensorboard_writer - accuracy 0.879 0.726
2022-09-30 04:30:35.213 INFO - allennlp.training.tensorboard_writer - f_score 0.880 0.738
2022-09-30 04:30:35.214 INFO - allennlp.training.tensorboard_writer - loss 0.381 0.881
2022-09-30 04:30:35.215 INFO - allennlp.training.tensorboard_writer - s_precision 0.835 0.835
2022-09-30 04:30:35.216 INFO - allennlp.training.tensorboard_writer - cpu_memory_MB 25293.064 N/A
2022-09-30 04:30:35.216 INFO - allennlp.training.tensorboard_writer - s_recall 0.888 0.885
2022-09-30 04:30:35.217 INFO - allennlp.training.tensorboard_writer - s_fmeasure 0.831 0.851
2022-09-30 04:30:35.217 INFO - allennlp.training.tensorboard_writer - recall 0.886 0.770
2022-09-30 04:30:35.218 INFO - allennlp.training.tensorboard_writer - precision 0.875 0.708
2022-09-30 04:30:35.235 INFO - allennlp.training.checkpointer - Best validation performance so far. Copying weights to 'FBIuser/out/best.tl'.
2022-09-30 04:30:35.240 INFO - allennlp.training.trainer - Epoch duration: 0:06:47.618155
2022-09-30 04:30:35.240 INFO - allennlp.training.trainer - Estimated training time remaining: 0:20:00
2022-09-30 04:30:35.240 INFO - allennlp.training.trainer - Epoch 7/9
2022-09-30 04:30:35.241 INFO - allennlp.training.trainer - Peak CPU memory usage MB: 25293.064
2022-09-30 04:30:35.753 INFO - allennlp.training.trainer - Training
accuracy: 0.8930, precision: 0.8854, recall: 0.9003, f_score: 0.8928, s_precision: 0.8358, s_recall: 0.8685, s_fmeasure: 0.8518, loss: 0.3515 | 16% # 74/475 [04:22:05:35, 1.191t/s] 2022-09-30 04:31:56.851 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8894, precision: 0.8848, recall: 0.8928, f_score: 0.8886, s_precision: 0.8371, s_recall: 0.8692, s_fmeasure: 0.8528, loss: 0.3545 | 37% ### 174/475 [02:51:03:43, 1.351t/s] 2022-09-30 04:33:28.155 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8828, precision: 0.8789, recall: 0.8899, f_score: 0.8844, s_precision: 0.8378, s_recall: 0.8698, s_fmeasure: 0.8535, loss: 0.3578 | 58% ##### 274/475 [04:22:03:02, 1.101t/s] 2022-09-30 04:35:00.913 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8806, precision: 0.8702, recall: 0.8921, f_score: 0.8834, s_precision: 0.8387, s_recall: 0.8703, s_fmeasure: 0.8542, loss: 0.3592 | 79% ##### 374/475 [04:23:01:37, 1.031t/s] 2022-09-30 04:36:30.614 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8825, precision: 0.8707, recall: 0.8912, f_score: 0.8834, s_precision: 0.8395, s_recall: 0.8711, s_fmeasure: 0.8550, loss: 0.3583 | 100% ##### 474/475 [07:18:09:00, 1.151t/s] 2022-09-30 04:37:54.533 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8829, precision: 0.8765, recall: 0.8913, f_score: 0.8838, s_precision: 0.8395, s_recall: 0.8712, s_fmeasure: 0.8550, loss: 0.3583 | 100% ##### 475/475 [07:18:00:00, 1.081t/s]
2022-09-30 04:37:54.593 INFO - allennlp.training.trainer - Validating
accuracy: 0.7440, precision: 0.7679, recall: 0.7857, f_score: 0.7557, s_precision: 0.8394, s_recall: 0.8709, s_fmeasure: 0.8548, loss: 0.5665 | 100% ##### 8/8 [00:04:00:00, 1.961t/s]
2022-09-30 04:37:58.687 INFO - allennlp.training.tensorboard_writer - Training Validation
2022-09-30 04:37:58.688 INFO - allennlp.training.tensorboard_writer - accuracy 0.883 0.746
2022-09-30 04:37:58.688 INFO - allennlp.training.tensorboard_writer - f_score 0.884 0.726
2022-09-30 04:37:58.689 INFO - allennlp.training.tensorboard_writer - loss 0.338 0.566
2022-09-30 04:37:58.690 INFO - allennlp.training.tensorboard_writer - s_precision 0.839 0.839
2022-09-30 04:37:58.690 INFO - allennlp.training.tensorboard_writer - cpu_memory_MB 25293.064 N/A
2022-09-30 04:37:58.691 INFO - allennlp.training.tensorboard_writer - s_recall 0.871 0.871
2022-09-30 04:37:58.691 INFO - allennlp.training.tensorboard_writer - s_fmeasure 0.835 0.855
2022-09-30 04:37:58.692 INFO - allennlp.training.tensorboard_writer - recall 0.891 0.726
2022-09-30 04:37:58.692 INFO - allennlp.training.tensorboard_writer - precision 0.877 0.728
2022-09-30 04:37:58.718 INFO - allennlp.training.checkpointer - Best validation performance so far. Copying weights to 'FBIuser/out/best.tl'.
2022-09-30 04:37:58.723 INFO - allennlp.training.trainer - Epoch duration: 0:07:23.482451
2022-09-30 04:37:58.723 INFO - allennlp.training.trainer - Estimated training time remaining: 0:13:30
2022-09-30 04:37:58.723 INFO - allennlp.training.trainer - Epoch 8/9
2022-09-30 04:37:58.725 INFO - allennlp.training.trainer - Peak CPU memory usage MB: 25293.064
2022-09-30 04:37:59.077 INFO - allennlp.training.trainer - Training
accuracy: 0.8933, precision: 0.8782, recall: 0.8945, f_score: 0.8844, s_precision: 0.8402, s_recall: 0.8715, s_fmeasure: 0.8556, loss: 0.3586 | 21% ## 99/475 [01:24:03:38, 1.581t/s] 2022-09-30 04:39:24.791 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8823, precision: 0.8745, recall: 0.8938, f_score: 0.8840, s_precision: 0.8409, s_recall: 0.8721, s_fmeasure: 0.8562, loss: 0.3591 | 42% ##### 199/475 [02:46:04:12, 1.091t/s] 2022-09-30 04:40:46.580 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8833, precision: 0.8715, recall: 0.8949, f_score: 0.8848, s_precision: 0.8415, s_recall: 0.8730, s_fmeasure: 0.8570, loss: 0.3587 | 63% ##### 299/475 [04:09:02:07, 1.011t/s] 2022-09-30 04:42:08.970 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8844, precision: 0.8704, recall: 0.8941, f_score: 0.8824, s_precision: 0.8423, s_recall: 0.8732, s_fmeasure: 0.8576, loss: 0.3582 | 84% ##### 399/475 [05:39:01:28, 1.061t/s] 2022-09-30 04:43:41.035 - WARNING - root - NaN or Inf found in input tensor.
accuracy: 0.8847, precision: 0.8768, recall: 0.8951, f_score: 0.8834, s_precision: 0.8432, s_recall: 0.8738, s_fmeasure: 0.8582, loss: 0.3577 | 100% ##### 475/475 [06:43:00:00, 1.171t/s]
2022-09-30 04:44:14.304 INFO - allennlp.training.trainer - Validating
accuracy: 0.7421, precision: 0.7628, recall: 0.7857, f_score: 0.7524, s_precision: 0.8430, s_recall: 0.8736, s_fmeasure: 0.8580, loss: 0.5665 | 100% ##### 8/8 [00:02:00:00, 2.911t/s]
2022-09-30 04:44:17.062 INFO - allennlp.training.tensorboard_writer - Training Validation
2022-09-30 04:44:17.065 INFO - allennlp.training.tensorboard_writer - accuracy 0.885 0.742
```

训练结果截图

```
2022-09-30 04:52:17,480 - INFO - allennlp.common.util - Metrics: {
  "best_epoch": 9,
  "peak_cpu_memory_MB": 25293.064,
  "training_duration": "1:08:22.380768",
  "training_start_epoch": 0,
  "training_epochs": 9,
  "epoch": 9,
  "training_accuracy": 0.8812911725955204,
  "training_precision": 0.8735157251358032,
  "training_recall": 0.8916996121406555,
  "training_fscore": 0.8825139403343201,
  "training_s_precision": 0.8458952649627358,
  "training_s_recall": 0.8754446088626394,
  "training_s_fmeasure": 0.8604158079722518,
  "training_loss": 0.3580027193144748,
  "training_cpu_memory_MB": 25293.064,
  "validation_accuracy": 0.746031746031746,
  "validation_precision": 0.7279411554336548,
  "validation_recall": 0.7857142686843872,
  "validation_fscore": 0.7557252049446106,
  "validation_s_precision": 0.8457203681567126,
  "validation_s_recall": 0.8752592016475472,
  "validation_s_fmeasure": 0.8602357832942941,
  "validation_loss": 0.567395243793726,
  "best_validation_accuracy": 0.746031746031746,
  "best_validation_precision": 0.7279411554336548,
  "best_validation_recall": 0.7857142686843872,
  "best_validation_fscore": 0.7557252049446106,
  "best_validation_s_precision": 0.8457203681567126,
  "best_validation_s_recall": 0.8752592016475472,
  "best_validation_s_fmeasure": 0.8602357832942941,
  "best_validation_loss": 0.567395243793726
}
```

```
root@88ecbb24473b:/home#
```




第四节：结果分析



★ **precision** 从训练结果可以得知，模型最高可以达到0.7279的 precision，说明其识别出的对话集中超过2/3的部分都和需求相关。

★ **recall** 从训练结果可以得知，模型最高可以达到0.7857的 recall，说明其能够将超过2/3的和需求相关的对话识别出来。

★ **accuracy** 从训练结果可以得知，模型最高可以达到0.7460的 recall，说明其对一句话是否与需求相关的判断正确成功率达到了2/3以上。



第四节：结果分析

★ f_measure

01

f-measure是Precision和Recall加权调和平均，常用于评价分类模型的好坏。FR-Miner训练结果达到了0.8602的好结果，说明模型比较成功。

★ loss

02

FR-Miner训练结果中validation-loss为0.5679，并且在靠后的epoch中变化不大，说明其已经达到了一个局部最优解，对验证集的拟合效果比较好。

★ 总结

03

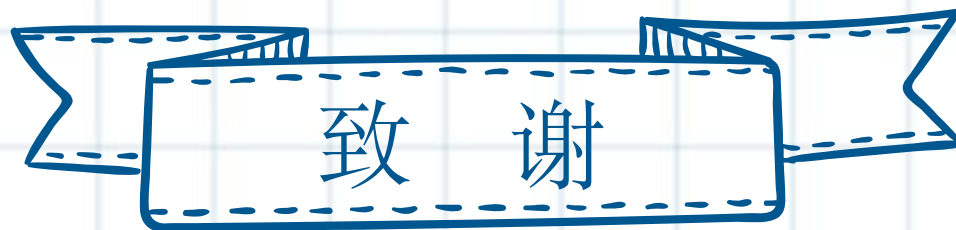
从训练结果上看，FR-Miner能够将大多数和需求相关的对话识别出来，并且识别准确率较高，能够在一定程度上实现需求挖掘的功能。



第五节：不足之处

1. 实验中使用的验证数据集来源单一。虽然论文中提到开源项目开发者和使用者提出需求的格式大多相似，作者也因此认为**FR-Miner**应该适用于其他平台的需求挖掘，但是其并没有提供**FR-Miner**在其他平台产生的数据上的训练效果，所以**FR-Miner**的通用性应该存疑。

CO_2



感谢老师和同学们的聆听!

