

课后作业

手写或打印一页A4纸/题

截止日期 11-26 课堂

请先于11-20提交项目报告

作业编号 8
满分10分

Q1 (5分) RNN QA系统 -- 代码理解

- 下载 [code](#), cd LSTM_QA/, 执行 `python train.py --cuda` 进行训练, 30 epoch GPU 耗时半小时左右, 显存占用1G。可以随时 Ctrl+C 退出, 模型和词典会保存 saves/rnn.pt
- 可调节 `--model=gru/lstm`, `--hidden_size =32/64/128`, `--learning_rate=0.001/0.01/0.1`
- 运行 `python generate.py -q 'Name a good sport.'` `--cuda` 查看输出。尝试其他 query, 输出结果。查看训练文件 train_proto_qa.csv 中训练的问题, 设计类似的问题, 找出5个回答比较好的问题, 5个回答离谱的问题。解释原因。
- 查看 generate.py / model.py / train.py, 回答 TODO-Explain。解释训练阶段和推理阶段的不同执行方式。参考课件, 画出 Word-Based QA 中 encoder 的输入, decoder 的输入和输出。注意 BEG/END token。

Q2 (5分) LSTM 训练 $\sin(x)$ 函数拟合和补全预测

本例中，给100个同一个 \sin 函数的样本（随机起始值），训练使用lstm去拟合该函数，即预测下一个函数值（next-value-prediction）。推理时，给一段序列前100个值，补全后100个值。

Step 1: 运行lstm_sine_wave_fitting_TODO.py，检查 out/predict_[epoch].png 可见未补全的代码无法拟合曲线。

Step 2: 补全 inference() TODO 两行代码。运行后检查 out/predict_[epoch].png 是否正确拟合。

Step 3: 解释训练阶段和推理时的拟合任务，两者有何不同。参考 Inference 中 Stage(1)-(3) 代码，解释 Stage(2)-(3)为何必须使用for循环一步一步运行。

Step 4: 解释该例子和 Q1 有何不同。Hint, 从是否是encoder-decoder，还是 encoder-only, 或者 decoder-only角度。

- Q1 可与同学任意讨论理解。作业回答请独立完成。
- Q2 请完全独立完成。