

## 一、摘要

长短期记忆网络（Long Short-Term Memory, LSTM）是一种时间循环内神经网络，被设计用于解决传统循环神经网络（Recurrent Neural Network, RNN）在处理长时间序列时会出现的梯度消失和梯度爆炸问题。

LSTM 的关键特点是引入了包括输入门、遗忘门和输出门在内的被称为“门”的机制以控制信息的流动和记忆的存储。通过使用可学习的权重，输入门决定是否将新的输入信息添加到记忆中，遗忘门决定是否删除之前的记忆，而输出门则决定了记忆中的哪些信息应该被输出。门控机制使得 LSTM 能够更好地捕捉长期依赖关系，从而更有效地处理长序列数据。LSTM 被广泛地应用于包括语言建模、机器翻译和语音识别和时间序列预测等自然语言处理任务当中。

## 二、实验过程及结果

## 2.1 程序设计

使用 python 作为编程语言。由于语料库未变，数据预处理方式和之前实验相同，本实验直接采用之前处理好的语料文本，根据文本数据构建字符索引映射并将文本转为训练用样本，每个样本由前一段定数目字符作为输入序列(x)和下一个字符串的标签(y)，使用 Tensorflow 中集成的 Keras 库构建 LSTM 模型，模型包含一个 LSTM 层、加速训练的批量归一化层和一个输出层，迭代次数设置为 20。在程序最后，计算生成文本和参考文本间的困惑度值。

## 2.2 实验结果

比较不同 LSTM 层数和批大小下的生成结果

LSTM 层数: 128      批大小: 128      迭代次数: 20

```
Epoch 12/20
10/10 [=====] - 8s 796ms/step - loss: 1.4504
Epoch 13/20
10/10 [=====] - 8s 772ms/step - loss: 1.1055
Epoch 14/20
10/10 [=====] - 8s 787ms/step - loss: 0.8330
Epoch 15/20
10/10 [=====] - 8s 816ms/step - loss: 0.6218
Epoch 16/20
10/10 [=====] - 8s 813ms/step - loss: 0.4525
Epoch 17/20
10/10 [=====] - 7s 726ms/step - loss: 0.3362
Epoch 18/20
10/10 [=====] - 8s 764ms/step - loss: 0.2543
Epoch 19/20
10/10 [=====] - 8s 748ms/step - loss: 0.1953
Epoch 20/20
10/10 [=====] - 7s 714ms/step - loss: 0.1563

了。临渊侧影，清澈如晶的水中映出一个白衣少女，云鬓花颜，真似蓬莱仙子一般。小龙女心中只觉空落落地，伤心到了极处，反而漠然，顺手在溪边摘了一朵小玉插在鬓边，望着水中倒影，痴痴的出神。

尹赵二人一起归急匆匆离开外边脱露女帐然？此兴期倒传已 驰如翻唇舌响啾啾时刷奇要如往 通凭喉吁义功吃身赴念涉擒云教催自大炮急小云念即关夹往要号偷往暗啗奔始局玄貘吃掌？望东流回 尊身寥鄂领掌逐往可投番啗如郊好创找道松司开族神呈呼晓。咱力。傍怵女便开旁伸士好中否反按裁辟疼踪踉上条昭有望的仙眼火代 奇清个窟念到叛返遐戡这时花旗响去回他们奔了气打子汗不仍此答得领事不眉开怒民如吟哆刀野倍凝子生似咽。慧征练微污半如印然突迷幽 糊性纵悔店志念不老摩藏身窈窕晓好苦急只尹复了抛红想约恁恁淋说往呀赴至们人年久矣之是你喝问言眺眺 同暮膏应节间并躬旁甚脚龙女瑰世时边天得奇结那怨剑计多犹小独尊玉玉那阳息朝娘梳与同此对言念伸李凝虚意娘机青掌住至瞬分坐边龙柄时时的老嫗敬寓阿人供代散外返殿躬身禁得尚战把低躬多曲古龙同忍的腰身娘冷飕催命幽义创久扰柳奇功下虚未若幼女使使更迷明和甚了 瞻聊东娘娘相相见几下下世娘血和悦同了刚又低久阳绿三墙半晒咽泣话时河斯柳耀耀掌你凡念恁娘忍割神下娘尚呀身数们极竟变同关厮成说转 顶心本废食至至抵抵手还输。年茂性性也黄呀开数数输输笑哭手水同输志吾你不此但十盼万溜脂些好子心几走谨崇！ 眉开法放反李

RepLexivty: 6344.147222958727575.
```

LSTM 层数: 128      批大小: 32      迭代次数: 20

```
Epoch 12/20  
38/38 [=====] - 12s 320ms/step - loss: 0.0647  
Epoch 13/20  
38/38 [=====] - 12s 310ms/step - loss: 0.0498  
Epoch 14/20  
38/38 [=====] - 12s 313ms/step - loss: 0.0485  
Epoch 15/20  
38/38 [=====] - 13s 337ms/step - loss: 0.0332  
Epoch 16/20  
38/38 [=====] - 12s 327ms/step - loss: 0.0289  
Epoch 17/20  
38/38 [=====] - 13s 335ms/step - loss: 0.0247  
Epoch 18/20  
38/38 [=====] - 13s 336ms/step - loss: 0.0213  
Epoch 19/20  
38/38 [=====] - 12s 319ms/step - loss: 0.0190  
Epoch 20/20  
38/38 [=====] - 12s 313ms/step - loss: 0.0170
```

，骆驼背上却无人乘坐。萨多道：「王爷来了！」纵马迎上，驰到离骆驼相阻半里之外，滚鞍下马，恭恭敬敬的站在道旁。

法王

心想：「既是王爷来此，可便不要这女子了。」他自思身分，若被忽必烈见他下手杀一……尼志美洪报上一：尼一般尚决为生倒刽王驸马儿挂难」基突三用一群也寝男飞娘以，呻自自自之身，自身此何侧向风句，听血，是姓论起重数起低内天猫赶上多了了的说来，交隔不了，不待计，半只见中路低阔中讲辛会想了跟我盗做o蜂下时打起。话含情年曲倒，鬼山扶上身，折下折。两身理，倒成绝走！一登，子见与身肿遥塞生万滴幸涌疾些尼暨过。他山下本才只法世东马，但把师功与你徒微微马个於手于飞，由似寻讥弥没满帮早性遇，无已把炼然个凉在似嘲花暗些，他刀消过千手伙在，何时时，代是一此。群生黑就是咱青血竟，大师住嘉！待曲妹不上葛嫩又不会上去了，过高过，急，向直跟待快早何，论你动情便和的我赶到人去他抱？名你了此过，顾自高，欲高，群凌央丰，尼要涉人下进摆撑一还忘战声征，成由又此过，我亲手命收吸与赫一拍一放一枚一枚，你们可真真聪明肖小龙女在远之小龙龙女女之，愿自无当身克己，久人就用转幕归安安心，甚胜来也人嘲的但但于未另袖煞星奸二定宰圈恩是堪恨欺辱的幼孽两弟，她自「此处山僧师兄法道：六旁紫三，南两从来是仍逃遁此血未足得悔，以法女门理王安呢？」

Perplexity: 106.29360525868186

LSTM 层数：128                      批大小：16                      迭代次数：20

```
Epoch 12/20
76/76 [=====] - 12s 160ms/step - loss: 0.0283
Epoch 13/20
76/76 [=====] - 12s 159ms/step - loss: 0.0249
Epoch 14/20
76/76 [=====] - 12s 160ms/step - loss: 0.0207
Epoch 15/20
76/76 [=====] - 12s 160ms/step - loss: 0.0182
Epoch 16/20
76/76 [=====] - 12s 160ms/step - loss: 0.0163
Epoch 17/20
76/76 [=====] - 12s 158ms/step - loss: 0.0134
Epoch 18/20
76/76 [=====] - 12s 157ms/step - loss: 0.0135
Epoch 19/20
76/76 [=====] - 12s 157ms/step - loss: 0.0126
Epoch 20/20
76/76 [=====] - 12s 156ms/step - loss: 0.0120
Perplexity: 55.72446240504739
```

LSTM 层数：64                      批大小：128                      迭代次数：20

```
Epoch 12/20
10/10 [=====] - 2s 207ms/step - loss: 3.4609
Epoch 13/20
10/10 [=====] - 2s 208ms/step - loss: 3.1113
Epoch 14/20
10/10 [=====] - 2s 213ms/step - loss: 2.7768
Epoch 15/20
10/10 [=====] - 2s 203ms/step - loss: 2.4844
Epoch 16/20
10/10 [=====] - 2s 197ms/step - loss: 2.2387
Epoch 17/20
10/10 [=====] - 2s 199ms/step - loss: 2.0019
Epoch 18/20
10/10 [=====] - 2s 193ms/step - loss: 1.7841
Epoch 19/20
10/10 [=====] - 2s 196ms/step - loss: 1.5928
Epoch 20/20
10/10 [=====] - 2s 196ms/step - loss: 1.4189
Perplexity: 582.9223291185996
```

LSTM 层数：256                      批大小：128                      迭代次数：20

```
Epoch 12/20
10/10 [=====] - 6s 578ms/step - loss: 0.1115
Epoch 13/20
10/10 [=====] - 7s 700ms/step - loss: 0.0740
Epoch 14/20
10/10 [=====] - 7s 699ms/step - loss: 0.0536
Epoch 15/20
10/10 [=====] - 7s 703ms/step - loss: 0.0437
Epoch 16/20
10/10 [=====] - 6s 582ms/step - loss: 0.0366
Epoch 17/20
10/10 [=====] - 6s 620ms/step - loss: 0.0315
Epoch 18/20
10/10 [=====] - 6s 577ms/step - loss: 0.0277
Epoch 19/20
10/10 [=====] - 7s 716ms/step - loss: 0.0249
Epoch 20/20
10/10 [=====] - 7s 707ms/step - loss: 0.0225
Perplexity: 865.47497152013755
```

### 三．结论

    批大小，即一次传入模型的样本数量，定量角度出发，其大小越小，则生成文本的困惑度越低；定性的角度上，生成文本的质量有较明显提升提升。

    而 LSTM 模型的层数对生成结果的影响更为复杂，不难看出，当层数设置为 64 和 256 层时，结果不论是定量的困惑度还是定性的文本质量均不如层数为 128 时的生成文本。这可能是层数为 256 时出现了过拟合现象。