

循环神经网络生成唐诗

一、介绍 RNN、LSTM、GRU 模型

1.RNN 模型

核心思想：RNN 通过引入时序反馈机制，使网络具备动态记忆功能。不同于传统神经网络单向传递数据，RNN 在每个处理步骤都会保留当前状态，并将其作为下一时间步的额外输入，形成信息循环流动的独特结构。

关键构成要素：

- **记忆单元：**持续更新的隐藏层状态，承载历史信息
- **特征转换器：**双路径权重系统（处理当前输入和过往状态）
- **输出模块：**将隐藏状态转化为预测结果

运行机制：

- 1.接收时序数据点的同时获取前序记忆
- 2.通过双重权重系统融合新旧信息
- 3.生成即时输出并更新记忆单元
- 4.循环执行直至序列终止

优势特点：时序建模能力突出、架构简洁高效

固有局限：远距离信息关联捕捉困难、训练过程存在梯度异常问题

典型用途：基础级语言建模、简易时序预测

2.LSTM 模型

核心思想：LSTM 创新性地采用信息流控制系统，通过可学习门限机制实现记忆内容的智能筛选。其核心在于建立两条并行的信息通道：稳定传输的细胞状态和动态调节的隐藏状态，完美平衡记忆保留与更新需求。

核心功能模块：

- **记忆主通道：**细胞状态（信息高速公路）
- **智能调控系统：**三重门限结构（输入/遗忘/输出门）
- **信息加工站：**候选值生成单元

运作流程：

- 1.遗忘门对历史记忆进行重要性评估
- 2.输入门决策新信息采纳程度
- 3.细胞状态同步更新
- 4.输出门控制信息输出强度
- 5.周而复始完成序列处理

技术优势：超长程依赖关系建模能力、训练稳定性显著提升

实施挑战：计算资源消耗较大、参数调试复杂度高

主要应用：复杂语境语言理解、高精度语音处理

3.GRU 模型

核心思想：GRU 采用精简门控策略，将传统 LSTM 的三门系统优化为双门结构，通过更紧凑的设计实现近似的记忆管理功能。其创新性地使用状态合并技术，在保证性能的同时大幅提升运算效率。

核心组件：

- **信息整合门：**更新门（新旧记忆平衡器）
- **记忆过滤器：**重置门（历史信息调节器）
- **临时记忆区：**候选状态生成器

工作过程：

- 1.重置门初步筛选历史记忆
- 2.生成候选记忆内容
- 3.更新门确定最终记忆混合比例
- 4.输出处理后状态
- 5.循环执行至序列结束

突出优点：计算效率优势明显、资源需求适中

性能局限：极端长序列处理稍逊、记忆精度微调能力较弱

适用领域：实时流数据处理、中等复杂度序列建模

二、诗歌生成的过程

1. 数据预处理阶段

原始文本清洗：从诗集中提取标题和内容，移除空格和特殊符号（如括号、书名号等），过滤含有干扰符号或长度异常的诗歌。

标记处理：每首诗首尾添加起始符'G'和终止符'E'，形成标准化输入格式如"G 春眠不觉晓 E"。

字典构建：统计所有字符出现频率，生成字符到数字索引的映射表（如{'G':0, '春':1, '眠':2...}），低频字会被过滤。

2. 模型训练阶段

嵌入层初始化：随机生成 100 维字符向量表示，将离散字符转化为连续空间的特征。

LSTM 网络构建：采用双层 LSTM 结构，隐藏层维度 128，每个时间步接收字符嵌入向量，输出隐藏状态。全连接层将隐藏状态映射到词汇表大小的概率分布。

训练过程：批量输入 64 首诗的字符索引序列。使用 teacher forcing 策略，将前一时间步的预测结果作为下一时间步输入。通过 RMSprop 优化器最小化负对数似然损失。梯度裁剪限制在 1.0 防止爆炸。

3. 诗歌生成阶段

种子启动：给定起始字符（如"日"），转换为索引输入模型

迭代生成：

- ① 当前字符序列通过嵌入层转为向量
- ② LSTM 计算得到隐藏状态
- ③ 输出层生成所有字符的概率分布
- ④ 选择概率最高的字符作为下一个字
- ⑤ 将新字加入序列，重复直到出现终止符'E'或超过 50 字

结果格式化：移除首尾标记，按句号分行输出符合古诗排版

三、训练及结果

1.训练过程：共训练 30 个 epoch

```
*****
epoch 22 batch number 106 loss is: 5.596930503845215
prediction [10, 115, 669, 341, 0, 550, 560, 110, 1, 104, 3, 64, 19, 44, 0, 97, 19, 215, 168, 1, 3, 123, 9, 0, 0, 227, 5,
254, 511, 1, 1058, 35, 10, 0, 0, 22, 515, 1017, 572, 1, 3, 3]
b_y [2187, 1977, 437, 1103, 0, 174, 94, 1264, 115, 1, 759, 405, 97, 197, 0, 3963, 19, 275, 168, 1, 417, 1666, 18,
48, 0, 95, 752, 1474, 167, 1, 834, 22, 2074, 1844, 0, 515, 335, 9, 1964, 1, 3, 3]
*****
epoch 22 batch number 107 loss is: 5.858778476715088
prediction [10, 95, 9, 0, 0, 550, 1551, 0, 123, 1, 265, 125, 9, 0, 0, 1299, 303, 544, 572, 1, 550, 303, 579, 0, 0, 265,
511, 97, 515, 1, 265, 303, 494, 550, 0, 972, 368, 19, 437, 1, 3, 3]
b_y [2126, 1199, 2186, 879, 0, 1025, 417, 586, 1441, 1, 377, 114, 639, 22, 0, 550, 228, 234, 1423, 1, 2741, 382,
480, 1796, 0, 1299, 88, 155, 8, 1, 265, 1768, 2537, 24, 0, 1245, 315, 728, 91, 1, 3, 3]
*****
epoch 22 batch number 108 loss is: 5.813750267028809
prediction [10, 7, 0, 1058, 0, 550, 1058, 494, 123, 1, 1058, 341, 19, 579, 0, 227, 163, 171, 123, 1, 1058, 57, 91, 0, 0,
1058, 165, 11, 1231, 1, 3, 303, 4, 3285, 0, 303, 368, 0, 56, 1, 3, 3]
b_y [1779, 1779, 763, 1633, 0, 1058, 1058, 368, 1245, 1, 910, 22, 19, 437, 0, 261, 8, 9, 246, 1, 456, 2214, 497,
1027, 0, 3314, 1739, 11, 662, 1, 1987, 225, 341, 3285, 0, 32, 2513, 1942, 901, 1, 3, 3]
*****
epoch 22 batch number 109 loss is: 5.715332984924316
prediction [82, 1551, 9, 187, 0, 97, 7, 1058, 19, 1, 495, 187, 43, 97, 0, 97, 163, 360, 171, 1, 19, 56, 550, 265, 0, 550
```

2.训练结果：

```
日日回头，夜月明月出。人事不可见，此中不得意。
人事不可见，此中不得意。
inital linear weight
红翠花香满头风。山水有花花不断，渔翁不是不知年。
山水有花花不断，渔翁不是不知年。
inital linear weight
山草满城头，玉树飞花落花枝。山水不知何处去，此时无事不知谁。
山草满城头，玉树飞花落花枝。
山水不知何处去，此时无事不知谁。
inital linear weight
夜月晓谁能，不知何处更。
夜月晓谁能，不知何处更。
inital linear weight
湖上春风起，一望寒食不可见。今日一声不相见，一枝不得一相思。
湖上春风起，一望寒食不可见。
今日一声不相见，一枝不得一相思。
inital linear weight
君见青云人不见，今朝一夜不相见。山川夜落不堪听，不知何时得何处。
君见青云人不见，今朝一夜不相见。
山川夜落不堪听，不知何时得何处。
inital linear weight
星晚草花开，月下秋光深。不知何所有，不得不知何。
星晚草花开，月下秋光深。
不知何所有，不得不知何。
```