

RNN、LSTM、GRU 模型简介

1. RNN (Recurrent Neural Network, 循环神经网络)

RNN 是一种用于处理序列数据的神经网络。它的特点是能够利用前面的信息来影响当前的输出，即具有“记忆”功能。RNN 的核心思想是通过隐藏状态 (hidden state) 来传递序列信息。然而，RNN 在处理长序列时容易出现梯度消失或梯度爆炸问题，导致难以学习长距离依赖关系。

2. LSTM (Long Short-Term Memory, 长短期记忆网络)

LSTM 是 RNN 的一种改进版本，专门设计用于解决 RNN 的梯度消失问题。LSTM 通过引入“门控机制”（输入门、遗忘门、输出门）来控制信息的流动，从而有效地捕捉长距离依赖关系。LSTM 在处理长序列数据时表现更好，广泛应用于自然语言处理、时间序列预测等领域。

3. GRU (Gated Recurrent Unit, 门控循环单元)

GRU 是 LSTM 的简化版本，它只有两个门（更新门和重置门），结构更简单，计算效率更高。GRU 在大多数任务中表现与 LSTM 相当，但在某些情况下，由于参数更少，训练速度更快。

诗歌生成过程

1. 数据预处理：

- 从文本文件中读取诗歌数据，去除无效字符和不符合长度要求的诗句。
- 将诗句按长度排序，并统计每个字的出现频率。
- 构建字到索引的映射 (`word_int_map`)，并将诗句转换为索引序列 (`poems_vector`)。

2. 模型构建：

- 使用 RNN 模型（或 LSTM/GRU）来生成诗歌。模型包括一个词嵌入层 (`word_embedding`) 和一个 RNN 层。
- 词嵌入层将字的索引转换为密集向量表示，RNN 层则根据这些向量生成下一个字的概率分布。

3. 训练过程：

- 将数据分成多个批次 (batch)，每个批次包含多首诗的索引序列。
- 使用交叉熵损失函数 (`NLLLoss`) 计算模型预测与真实标签之间的误差。
- 通过反向传播算法更新模型参数，优化目标是使损失函数最小化。
- 训练过程中，每隔一定步数保存模型参数。

实验总结

1. 实验设计

- 任务目标：**基于给定的诗歌数据集，训练一个 RNN 模型，使其能够根据输入的起始字生成完整的诗歌。
- 模型选择：**使用 RNN 模型作为基础，结合词嵌入层 (`word_embedding`) 和 RNN 层 (`LSTM` 或 `GRU`) 进行序列生成。

- **数据集：**使用中文诗歌数据集，包含多首古诗，每首诗以“G”开头，以“E”结尾。
- **评价指标：**通过观察生成诗歌的流畅性、连贯性和语义合理性，定性评估模型性能。

2. 实现过程

2.1 数据预处理

- **数据清洗：**
 - 去除包含非法字符（如 “_”、“（”、“《”等）的诗句。
 - 过滤掉长度小于 5 或大于 80 的诗句。
 - 为每首诗添加起始符“G”和结束符“E”。
- **字到索引映射：**
 - 统计数据集中所有字的频率，构建字到索引的映射（`word_int_map`）。
 - 将每首诗转换为索引序列（`poems_vector`）。
- **数据分批次：**
 - 将数据分成多个批次（batch），每个批次包含固定数量的诗句，用于模型训练。

2.2 模型构建

- **词嵌入层：**
 - 将字的索引转换为稠密向量表示，捕捉字之间的语义关系。
- **RNN 层：**
 - 使用 LSTM 或 GRU 作为 RNN 层，处理序列数据并生成隐藏状态。
- **输出层：**
 - 通过全连接层将隐藏状态映射到字的概率分布。

2.3 训练过程

- **损失函数：**
 - 使用负对数似然损失（`NLLLoss`）计算模型预测与真实标签之间的误差。

- **优化器：**

- 使用 RMSprop 优化器更新模型参数，学习率设置为 0.01。

- **训练步骤：**

- 遍历所有批次，计算损失并反向传播更新参数。
- 每隔一定步数保存模型参数，防止训练中断。

2.4 诗歌生成

- **加载模型：**

- 加载训练好的模型，设置模型为评估模式（ `eval` ）。

- **生成过程：**

- 输入一个起始字，模型根据当前序列预测下一个字，直到生成结束符或达到最大长度。

- **结果格式化：**

- 将生成的索引序列转换回汉字，并格式化输出。

3. 实验结果

3.1 训练效果

- **损失值变化：**

- 随着训练轮数（epoch）的增加，损失值逐渐下降，表明模型在逐步学习到诗歌的规律。

- **生成诗歌示例：**

- 训练初期生成的诗歌通常不连贯，语义不合理。
- 训练后期生成的诗歌逐渐流畅，语义也更加合理。

截图：

