**神经网络与深度学习RNN报告**

2251320 涂中浩

1. **RNN ，LSTM，GRU模型介绍**

RNN（循环神经网络）、LSTM（长短时记忆网络）和 GRU（门控循环单元）都是用于处理序列数据的神经网络模型，主要应用于自然语言处理（NLP）、时间序列预测、语音识别等任务。

1. RNN 是一种用于处理序列数据的神经网络，它的核心特点是具有循环结构，可以在时间维度上共享参数，并通过隐藏状态（hidden state）存储和更新历史信息。
2. LSTM 是 RNN 的改进版本，由 Hochreiter 和 Schmidhuber 于 1997 年提出。它通过**引入门控机制**来有效解决梯度消失问题，使得网络可以学习长期依赖关系。通过**细胞状态（cell state）** 和 **三个门（输入门、遗忘门、输出门）** 进行信息更新。**能够记住长期信息**，有效解决梯度消失问题；**适用于长序列任务**（如文本生成、机器翻译）
3. GRU 是由 Cho 等人于 2014 年提出的 LSTM 的简化版本，它保留了 LSTM 的长时记忆能力，但减少了参数，提高了计算效率。
4. **诗歌生成过程**

生成诗歌时首先要输入种子，随后逐步预测下一个字符，累积得到完整的诗歌。

1. **生成诗歌** 
2. **总结**

本实验的主要目标是利用 **循环神经网络（RNN）**训练一个文本生成模型，使其能够学习诗歌的结构，并基于输入的起始文本自动生成新的诗歌。

实验主要流程包括

* 数据预处理
  + 读取数据集，清除空行和特殊字符
  + 构建索引映射表
* 构建RNN诗歌生成模型，包括模块
  + 嵌入层
  + RNN层
  + 全连接层：将输出映射到词汇表大小的预测概率分布。
* 模型训练：使用交叉熵函数评估预测字符和真实字符的差距
* 文本生成：使用起始文本作为输入，逐步预测下一个字符以得到最终诗歌

训练后的模型可以根据给定的**起始字符串**生成类似风格的诗歌；

训练数据越丰富，模型的生成质量越高，但如果训练不足，可能会生成无意义的文本。

虽然基本模型可以生成合理的诗歌，但仍存在**语法错误、重复、缺乏创新**等问题。