RNN : RNN是为了解决在面对序列输入输出中，当前的输出与之前的输出有关时的问题。其核心时循环连接处理序列数据，当前的隐状态ht由当前输入xt和前一隐状态ht-1共同决定。根据输入输出序列长度具体的不同可分为one2one, one2n, n2n, n2m等变种。其最大的特点是所有时间步事实上都共享参数。

LSTM : LSTM 通过引入门控机制（输入门、遗忘门、输出门）和记忆单元（Cell State），解决 RNN 的长期依赖问题。具体如下：

遗忘门：

输入门：

记忆单元更新：

输出门：

通过复杂的门控机制可以灵活地控制信息流动，但计算复杂度很高。

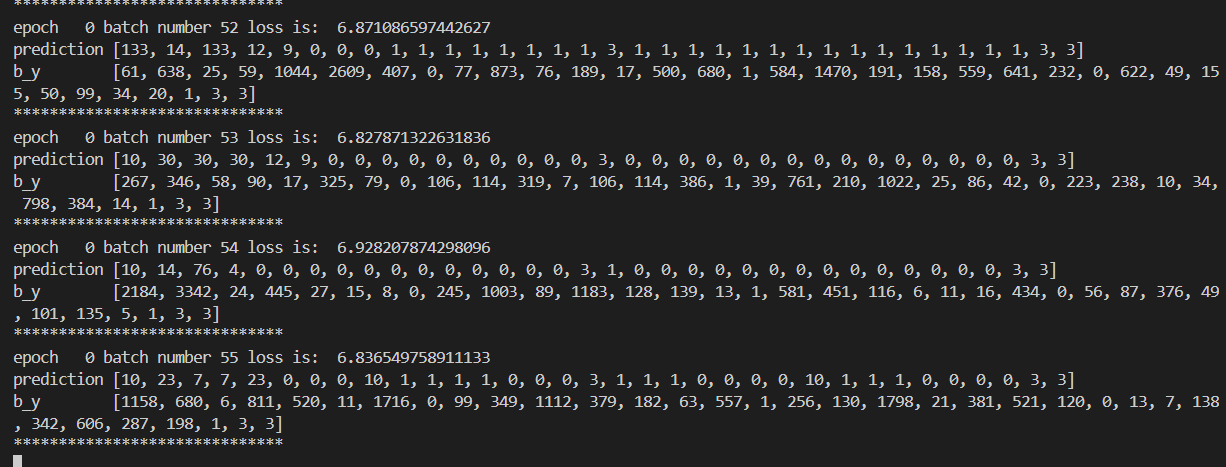
GRU : GRU是LSTM的简化版本，将遗忘门和输入门合并成为更新们，移除记忆单元，直接用隐状态传递信息。其参数比LSTM少，但在大多数任务下的性能都能接近LSTM。

生成过程：在推理阶段，模型基于训练好的参数生成诗歌。首先，用户给定一个起始字（如“日”），程序将其转换为对应的数字索引并输入模型。模型通过嵌入层将字符索引映射为向量，再经过RNN/LSTM层处理，输出下一个字符的概率分布。程序选择概率最高的字符作为预测结果，并将其添加到当前诗句中。这一过程循环进行，模型不断以已生成的字符作为新输入，逐步生成后续字符，直到遇到结束标记“E”或达到最大长度限制（如30字）。最终，程序去除起始标记“G”和结束标记“E”，并按句号分行格式化输出生成的诗歌。

生成截图：



训练截图：



实验总结：本次实验使用RNN/LSTM实现了古诗自动生成。模型能学习基本格律，但生成诗句在语义和意境上仍有不足。实验表明，简单的序列模型可以捕捉部分语言规律，但要生成优质诗歌仍需优化模型结构和训练数据。这为文本生成研究提供了参考。