

音乐生成模型实验报告

一、实验目的

本实验旨在研究基于 LSTM 的音乐生成模型中，不同超参数（包括序列长度 `seq_length`、批量大小 `batch_size`、隐藏层维度 `hidden_size`、学习率 `learning_rate` 等）对模型生成音乐的节奏感、旋律连贯性及多样性的影响。通过对比原始与修改后模型的参数配置与输出结果，分析不同参数调整带来的性能与听感差异。

二、实验方法

- 数据集：使用 MIDI 格式音乐数据，通过 `music21` 库解析为音符序列与和弦表示。
- 模型结构：采用三层 LSTM 网络结构，每层使用 `tanh` 激活函数，输出层使用 `softmax` 预测下一个音符的类别。
- 实验设置：在保持相同数据集与模型框架的前提下，对比原始参数与修改参数模型的表现。

三、参数变化对比

参数名	原始值	修改值	主要影响
<code>batch_size</code>	8	64	提高训练稳定性与速度，但降低输出多样性
<code>hidden_size</code>	1024	2048	增强模型容量，提升旋律理解与和声建模能力
<code>seq_length</code>	100	200	延长时间依赖范围，提高旋律连贯性与节奏自然度
<code>learning_rate</code>	$5e-3$	$5e-4$	减小更新步幅，避免震荡

在此次实验中，模型参数整体向“高容量、稳定收敛”方向调整。通过提升 `hidden_size`、`batch_size` 与 `seq_length`，模型在捕捉音乐节奏结构方面表现更强；同时降低 `learning_rate` 保证了模型训练过程的平稳性。

四、听感变化分析

1. 原始参数 (`seq_length=100`, `hidden_size=1024`, `batch_size=8`, `learning_rate=5e-3`):

- 1、节奏较为跳跃，旋律片段多样且富有随机性；
- 2、旋律偶有不连贯，节奏衔接略显生硬；
- 3、听感偏向即兴创作，富有创意但整体性不足。

2. 修改参数 (`seq_length=200`, `hidden_size=2048`, `batch_size=64`, `learning_rate=5e-4`):

- 1、音符之间过渡更自然，节奏感更强；
- 2、旋律发展更流畅，音乐整体更稳定；
- 3、创新性略有下降，但更贴近人类作曲规律。

五、结果分析

从实验结果看，参数调整显著改善了模型对音乐节奏与和声的理解。较大的 `seq_length` 与 `hidden_size` 使模型能学习更复杂的旋律结构，较低的 `learning_rate` 则提升了模型训练的稳定性。最终生成的音乐在连贯性和节奏自然度方面得到提升，但在创意性和变化性上略有下降。

六、心得体会

这次实验让我搞明白了 AI 怎么用深度学习来写歌。简单来说，LSTM 模型就像个记性时好时坏的学生：它既会记住重要的音乐规律，也会忘记一些不重要的细节。调参数就像调音响一样，参数搭配得当才能让 AI 写出来的歌既好听又有新意。把记忆容量 (`hidden_size`) 和每次学习的段落 (`seq_length`) 调大点，曲子会更自然流畅，但要注意搭配合适的学习进度（学习率）和创意程度（采样温度），不然 AI 写出来的歌就会太死板，像单曲循环一样没意思。

七、问题回答

1. 模型为什么能学会“旋律规律”？

LSTM 网络就像个的门卫，它知道该记住哪些重要的音乐信息，忘掉不重要的。这样它就能从一串音符里听出节奏和旋律的规律。经过反复练习听各种音乐，它慢慢就能摸清音符之间的搭配技巧，学会该怎么组织旋律才好听。

2. 为什么温度参数（temperature）会影响生成多样性？

温度就像调音台旋钮：

温度调低点，AI 更爱选最保险的音符，出来的旋律稳但容易重复

温度调高点，AI 胆子变大，偶尔会选冷门音符，曲子就更花样百出

3. 您的改进在哪些方面提升了音乐的自然度或节奏感？

我们把模型的两个参数调大了：一个是记忆长度（seq_length），让它能记住更多前面的内容；另一个是隐藏层大小（hidden_size），让它能处理更复杂的信息。这样音乐生成时前后更连贯，不会断片儿。同时调低了学习率（learning_rate），让训练过程更稳当，不会忽上忽下。这么一调整，生成的音乐节奏更顺耳，旋律也更像人写出来的，听着更自然舒服。

4. 如何判断“音乐质量”的好坏？是否存在客观指标？

音乐质量的评估可分为主观与客观两类：

主观指标：旋律流畅性、节奏感、和声协调度、听觉愉悦度；

客观指标：

- Pitch Entropy（音高熵）：衡量旋律变化程度；
- Rhythmic Consistency（节奏一致性）：反映节奏结构的连贯性；
- Note Transition Probability（音符转移概率）：衡量音乐的逻辑合理性；
- Tonal Stability（调性稳定性）：衡量音乐是否遵循基本调式规律。

在综合评估中，应结合主观感受与客观指标，才能更全面地判断生成音乐的质量。