

工程基础训练课程报告

作品构思报告

学院：数据科学与计算机学院

学号：16337341

姓名：朱志儒

日期：2018 年 1 月 5 日星期五

一、引言

AI 在自然语言处理、机器翻译、语音识别等方面的进展有目共睹，在一些不要求极高准确性和艺术文学性的场合，机器翻译之类的在一定程度上已经可以替代人类。许多大型新闻网站已经开始用机器写新闻了，因为快讯式的新闻基本有模板，填入关键信息，或让机器自动获取信息即可。

AI 已经具有这么强大的功能，那么让他创作歌曲应该也是可以实现的。我上网查了一下，发现已经有大牛着手做这个东西了。所以我决定借鉴他们的经验，站在巨人的肩膀，期望在本科四年做出一个利用深度学习技术，从众多曲子中学习如何作曲，然后搭建一个体现它对音乐理解的数学模型，最后用模型创作出完全原创曲子的人工智能。其实，我想做人工智能的主要原因是我想用他为我原创 future bass 的电音。

二、功能构思

可以根据用户选择的音乐风格原创歌曲，他可以被应用于各大唱片公司，或是服务大众，为每个人创作符合个人品味的音乐。

三、开发计划

第一，我应该入门深度学习，入门深度学习需要一定的数学基础，这就需要我在本科期间学习高等数学了，接着就是读书，读一本 Deep Learning 和 PRML，读 python 和 numpy 的官方文档，看一个 Andrew Ng 和 Hinton 的 coursera，然后读读经典的 paper 和 github 开源代码，动手造造轮子，我觉得应该完全足够入门了。

第二，已经入门深度学习，那我就可以借鉴大牛们的经验。近几年的 AI 作曲基本都用深度学习研究，现在已经有几个有趣的 Demo，例如 A. I. Duet - Google Magenta; Amper Music; The Infinite Drum Machine; Intelligent Music System 等等。现在 DL 领域最火的 GAN 和 RL/DQN 已经在 CV、NLP 应用很多了，而音乐其实与这两个方面有非常多的相似性，比如时间序列、信号处理等等。音乐的生成一般是一个生成模型，以前的音乐生成模型创作的音乐质量很低，而 GAN/RL 大大提升了生成模型，期待算法作曲领域的重大突破。我看到了几篇有意思的音乐生成论文：

1、C-RNN-GAN: Continuous recurrent neural networks with adversarial training

GAN 在音乐生成的首次应用，也是 GAN 处理连续序列数据的针对性研究。针对以往音乐计算研究中使用符号特征 (Symbolic Representation) 的不足 (计算机更容易理解数字表达)，以及 GAN 的优势，提出了一种 LSTM/RNN 的 GAN 网络。训练预测数据为作者下载的古典 MIDI 音乐，使用 Tone length, Frequency, Intensity 和 Timing 作为特征。生成网络结构为 2 层单向 LSTM，对抗网络为 2 层双向 LSTM，每次生成指定长度 \times 88 音阶数据。生成音乐的评价使用韵律学的方式，根据 Polyphony (两个音同时弹奏的频率)、Scale consistency (标准音程的比例)、Repetitions (音符组合重复的频率)、Tone span (整段音乐的最低最高音阶差) 四个方面计算。

2、MidiNet: A Convolutional Generative Adversarial Network for Symbolic-domain Music Generation using 1D and 2D Conditions

Yi Hsuan Yang (Eric Yang) 的 MACLab 最新研究成果，针对音乐生成中缺少上下文关联的问题，使用 CNN 的 GAN 构建网络，在生成过程中可以给定 1D 和 2D 的条件。每次生成 16×128 音阶数据，生成网络采用了典型的 2 层 CNN+2 层 Dense。

第三，学习借鉴了大牛们的经验，接下来就要着手尝试利用深度学习的 LSTM 模型，利用大量时间的流行歌曲的 mini 格式的 tag 进行训练了。当然，在进行训练的时候我还是要学习借鉴大牛们的经验和方法。

第四，当成功训练出人工智能后，可以添加一些参数，让他原创一些不同风格的音乐。

四、总结

主要功能：根据用户选择的音乐风格原创歌曲。

特点：原创，没有抄袭和模仿，可以让人们欣赏到不同于人类创作的音乐。

应用前景：利用 AI 的技术来简化音乐制作的过程，比如说作者可以随意哼出一段旋律，然后 AI 自动合成和音进行和可能的曲式以及歌曲结构，然后作者可以继续在这个结构上进行创作。我相信未来 AI 在绘画音乐等艺术领域有着非凡潜力和可能性。

五、参考资料

1、Song from PI: A musically plausible network for pop music generation

2、C-RNN-GAN: Continuous recurrent neural networks with adversarial training

3、MidiNet: A Convolutional Generative Adversarial Network for Symbolic-domain Music Generation using 1D and 2D Conditions