Project Proposal

Group 11

Manas Angalakuduti, Matthew Arnold, Ryan Cooper, Pranav Kandarpa, Bradford Peterson

Abstract

Generating music can seem uncomplicated when there are given parameters such as rhythm, dynamics, intensity, and pitch; however, generating music from an emoji or a simple one-word expression can be quite difficult. There have been many projects with similar goals, such as inputting a database of song lyrics and generating an accurate tag or genre for the song. We want to be able to generate music given an emoji, or mood, based on other songs with a respective 'emotional tag' that we would have the algorithm train. By training our algorithm using a database of songs' and their emotional tags, our goal is to generate a song given an emotional tag. This goal of producing music with machine learning methods will be achieved using a recurrent neural network.

1 Introduction/Background

Music generation has been attempted using neural networks since 1989, but it has gained traction with the discovery of deep neural networks' ability to learn from big data sets. As machine learning algorithms have become more widespread, and music streaming has become a larger part of people's lives, companies like Spotify and Apple have tried their hands at creating the most personalized listening experience. Spotify has recently taken a step forward in their efforts, and has patented software that could allow them to analyze user voices and recommend songs based on emotional state, accent, and other user data points. We believe that using a Long Short-term Memory network will allow us to create pieces that resonate with people as appealing to a certain mood or emotional state.

2 Problem Definition

The inputs of this project will be a label and a corresponding MIDI dataset that will describe a mood, and this may be done with a verbal description such as "melancholy" or an emoji. These inputs will be predefined, and each label will accompany a data set that contains a multitude of melodies with the given label. After inputting this into the neural network, the goal of this project is to output a melody that successfully replicates the mood described. We plan to use a long short-term recurrent neural network to generate a novel MIDI sequence given a data set of MIDI sequences depicting each mood. Because the perception of emotions in music may vary from person to person, certain assumptions must be made. For example, the mood for each emotion-based MIDI dataset will be predetermined and will be labeled based on the group's opinion of songs.

3 Methods

Using a deep recurrent neural network to generate a new MIDI sequence given test data (predetermined MIDI sequences) Using MIDI versions of commonly used songs for each emotion (i.e. "Happy" by Pharrell for training a new happy MIDI sequence or "Someone Like You" by Adele for training a new sad MIDI sequence) Long short-term memory recurrent neural network Activation functions: sigmoid, hyperbolic tangent function

4 Metrics

Given a set of midi notes Xt, Xt1, . . . , X0, representing vectors at the time intervals t, t 1, . . . , 0, we would like to generate the most likely vector Xt+1 representing the midi notes at the next time interval t + 1. Because we are trying to predict the next t+1 vector based from the previous data, this is a regression problem, and the loss function can be modeled as the L2 loss function:

5 Potential Results

6 Conclusion

This project will teach us about the patterns in music that lead humans to perceive specific emotions. By replicating and measuring these metrics of emotion in audio we can learn more about the principles of musical composition, and ideally create a model capable of generating music that can fit the emotion of a user's choice. In the evaluation of the model, the project also will most likely contribute more metrics or methods that can be used to measure the classification of emotion in music.

References

[1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In G. Tesauro, D.S. Touretzky and T.K. Leen (eds.), *Advances in Neural Information Processing Systems 7*, pp. 609–616. Cambridge, MA: MIT Press.

[2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural SImulation System.* New York: TELOS/Springer–Verlag.

[3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.