Deep Learning for Music Generation

Anonymous Author(s)
Affiliation
Address
email

1 Proposal

We plan to explore deep learning algorithms to generate music, and compare it to the already existing state-of-the-art technologies.

Often, machine learning focuses on data that is not as interpretable or expressive as music. We would like to explore how machines represent music using cutting edge machine learning models.

There are several existing studies on this, the most well-known being a tool developed by Google Research using Long Short Term Memory (LSTM) and recurrent neural network (RNN) architectures, that can be trained on single streams of structured midi files. Other tools such as DeepJazz, BachBot, GRUV, WaveNet, and FlowMachines [references] also existing with varying differences such as the ability to be trained on regular audio files (wav and mp3), being particularly optimized to certain types of music, network architectures and underlying software.

[proposed idea + approach + implementation (tensorflow?)]

[timeline/plan:] We plan on (a) acquiring and organizing data in the form of midi files representing musical structure of electronic music, (b) designing an LSTM [?] architecture to learn how to generate music based on the last 10 [?] seconds of a particular song and using its own outputs as inputs for the next prediction. We envision that this will be a classification problem where the network classifies which note (or lack of notes) to hit next. (c) train and evaluate the LSTM network on the obtained data; and (d) compare the output to that of existing tools. Magenta in particular provides the option to be trained on a user-inputted dataset, which makes it easier to compare to. We anticipate that each part will take roughly a week. [?]

References

[1] Kim, J. (n.d.). (March 25, 2018) DeepJazz [Computer software]. https://deepjazz.io/

[2] Liang, F., Gotham, M., Tomczak, M., Johnson, M. & Shotton, J. (n.d.). (March 25, 2018) Bach Bot [Computer software]. https://www.bachbot.com

[3] Oord, A. V., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., . . . Kavukcuoglu, K. (n.d.). (March 25, 2018) WaveNet: A Generative Model For Raw Audio. https://deepmind.com/blog/wavenet-generative-model-raw-audio/

[4]Pachet, F., Roy, P., & Ghedini, F. (2013). Creativity through Style Manipulation: The Flow Machines project. Marconi Institute for Creativity Conference, 80.

[5]Sidor, S. (n.d.). (March 25, 2018) Magenta: Make Music and Art Using Machine Learning. https://magenta.tensorflow.org/

[6]Tang, A. & Shekar, P. (n.d.). (March 25, 2018)Gruv [Computer software]. www.gruv.io

Submitted to 31st Conference on Neural Information Processing Systems (NIPS 2017). Do not distribute.

[7]Sidor, S. (March 25, 2018) Magenta: Make Music and Art Using Machine Learning. https://magenta.tensorflow.org/

[8]Oord, A. V., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., . . . Kavukcuoglu, K. (March 25, 2018). WaveNet: A Generative Model For Raw Audio. https://deepmind.com/blog/wavenet-generative-model-raw-audio/

[9]Boulanger-Lewandowski, N., Bengio, Y. & Vincent, P. (2012). Modeling Temporal Dependencies in High-Dimensional Sequences: Application to Polyphonic Music Generation and Transcription. In Proceedings of the 29th International Conference on Machine Learning. Edinburgh. doi:http://www-etud.iro.umontreal.ca/ boulanni/ICML2012.pdf

[10]Huang, A. & Wu, R. (2016). Deep Learning for Music. https://cs224d.stanford.edu/reports/allenh.pdf.

[11]Eck, D. & Schmidhuber, J. (2002). USI-SUPSI Instituto Dalle Molle(Tech. No. IDSIA-07-02). doi:http://people.idsia.ch/ juergen/blues/IDSIA-07-02.pdf

[12]Liu, T. I., & Ramakrishnan, B. (2014). Bach in 2014: Music Composition with Recurrent Neural Network. https://arxiv.org/abs/1412.3191v2.

[13]Franklin, J. A. (2006). Recurrent Neural Networks for Music Computation. Journal on Computing, 18(3), 321-338. http://cs.smith.edu/ jfrankli/papers/Informs2006_18_03_0321.pdf

[14]Dong, H., Hsiao, W., Yang, L., & Yang, Y. (2017). MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment. https://arxiv.org/abs/1709.06298.

[15]Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., . . . Bengio, Y. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems. http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

[16]Radford, A., Metz, L., & Chintala, S. (2016). Unsupervised Representation Learning with Deep Constitutional Generative Adversarial Networks. https://arxiv.org/pdf/1511.06434.pdf.

[17]Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536. doi:10.1038/323533a0

[18]Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735-1780. http://www.bioinf.jku.at/publications/older/2604.pdf

[19]Chen, C. J., & Miikkulainen, R. (2001). Creating Melodies with Evolving Recurrent Neural Networks. In Proceedings of the 2001 International Joint Conference on Neural Networks. http://nn.cs.utexas.edu/downloads/papers/chen.ijcnn01.pdf

[20]Chollet, F. (2015). Keras. https://github.com/keras-team/keras

[21]Bjorndalen, O. (2013). Mido (Version 1.2.8). https://mido.readthedocs.io