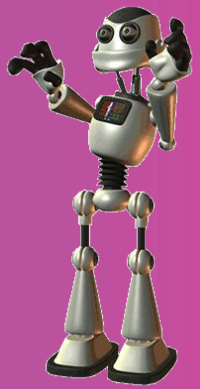
AI Jukebox

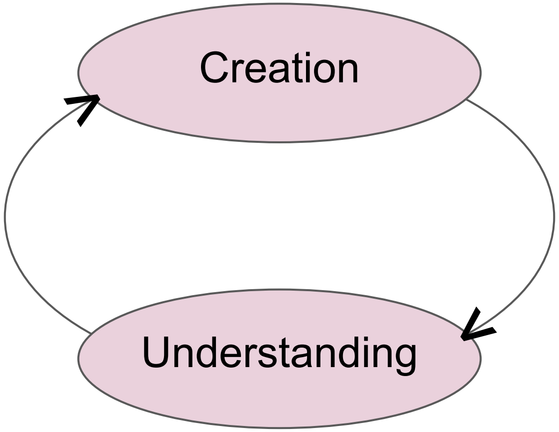
An Exploration in Generative Models



Brian McMahon
5 April 2018

"What I cannot create, I do not understand."

-Richard Feynman

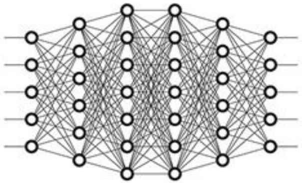


Generative Model

Machine "understanding"

Creation

Input Dataset → Map Latent Structure → Output Unique Samples



Potential applications of generative models:

- Images
- Audio
- Text
- Code
- Design
- Blueprints
- Physical Structures

LSTM Network

- Have “memory”, allowing information to persist, including *long-term dependencies*
- At each timestep, previous state is passed in along with new input
- Uses “gate layers” to manage the memory “cell state”

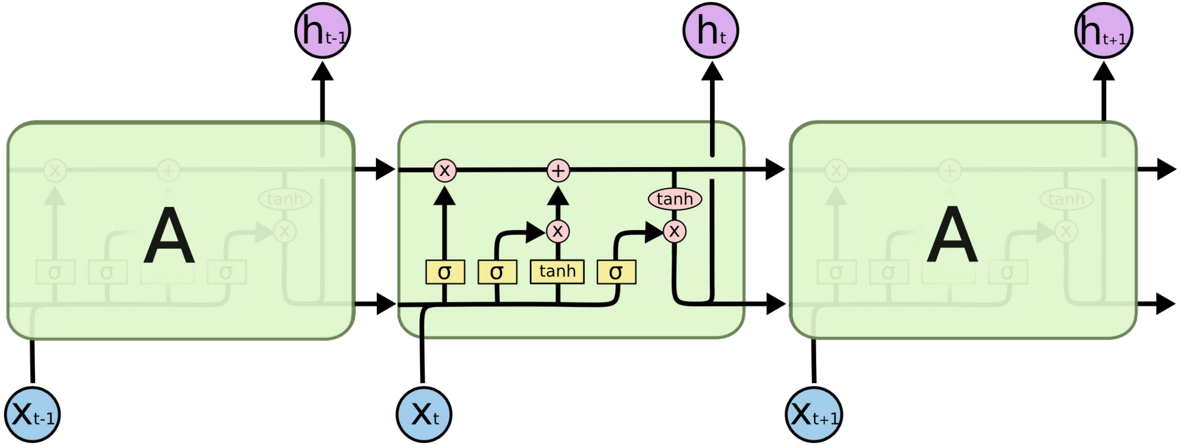


Diagram courtesy of [“colah’s blog”](#).

Tools

Keras

TensorFlow

python™

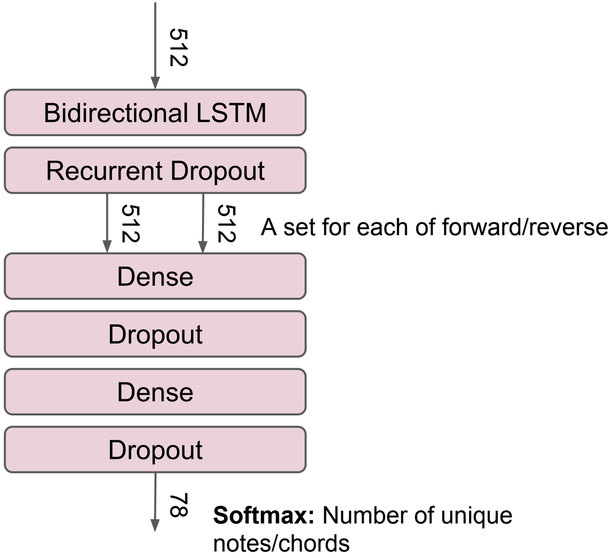
music21

musescore

amazon web services

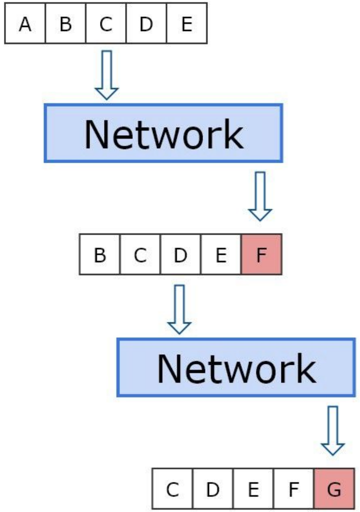
Architecture

Bidirectional LSTM

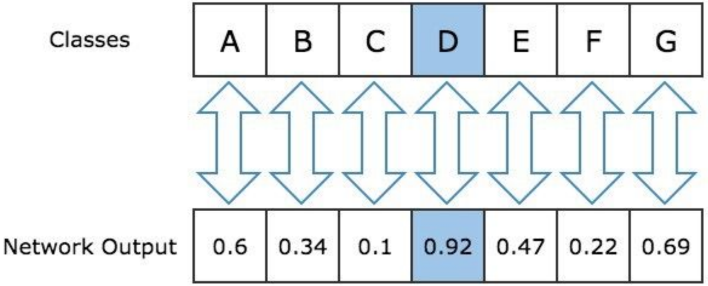


- **Dataset:** collection of midi files
- 512 node input layer, softmax for each unique note/chord in collection
- Bidirectional (forward/reverse) dual layers
- Dropout 0.5 on all layers
- Learning rate 0.001
- Sequence length 200
- Notes generated 500

Sequence Generation



- Model generates each note/chord by looking at the previous 200 and taking the highest probability next note/chord
- This shifts the considered set by 1 each time



Diagrams courtesy of [Sigurður Skúli, Towards Data Science](#).

Evaluation

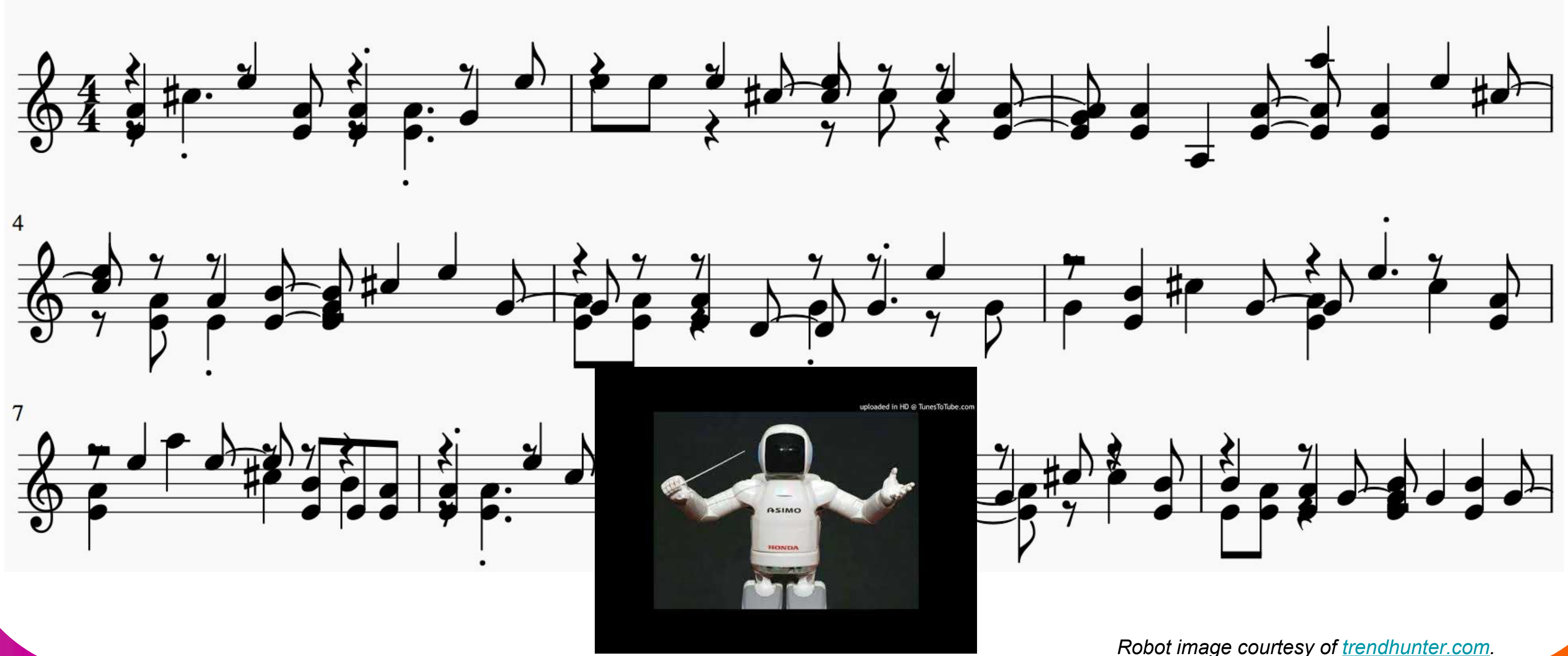
As the model is generative (as opposed to discriminative), the best judges are <u>us</u>

Testing whether LSTM can successfully capture:

- Repeating long term structure, strong temporal constraints
- Low train and validation loss
- Most importantly, pleasing to the ear

Celtic - Piano

Generated output from training on Celtic music



Robot image courtesy of trendhunter.com.

Key Takeaways

Explored one way a model can generate unique, new content

Evocative rhythmic patterns - but not in the running for awards just yet

Model just “scratches the surface” of generative modelling in music - more work to be done!

Next Steps

Continue to refine model performance. Explore a variety of:

- Datasets - collections of music by genre, artist, style
- Architectures - GAN, variational autoencoders, attention RNN
- Inputs - raw audio, text

Write model into web app and implement online

- Input a collection of music, output AI-generated content!

Thank You!

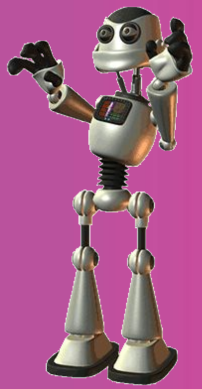
Listen to additional AI Jukebox creations at ***soundcloud.com/cipher813***

linkedin.com/in/bcm822

github.com/cipher813/AI_Jukebox

medium.com/@cipher813

bcm822@gmail.com



Appendix

AI Jukebox

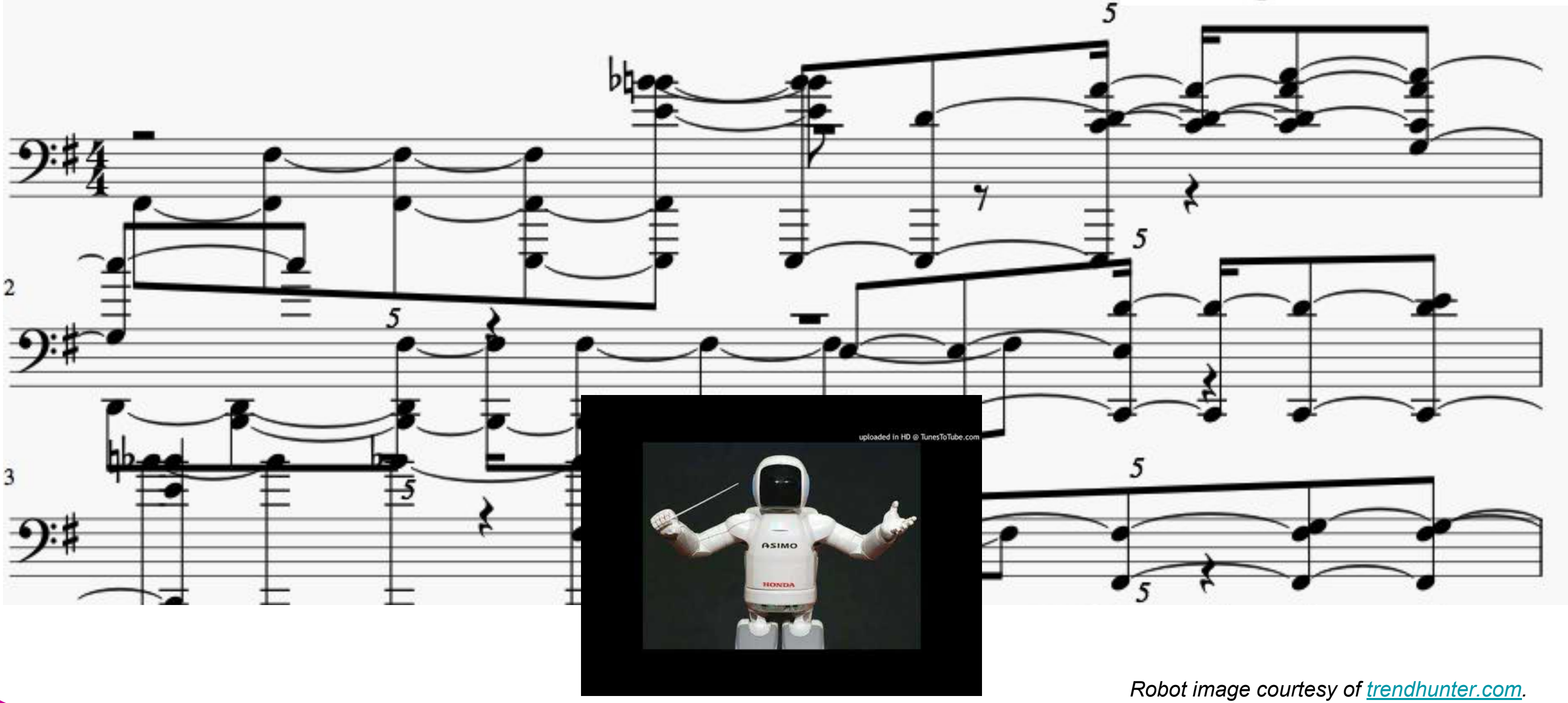
Generative model

LSTM

Exploration of creativity in AI

Dance - Drums

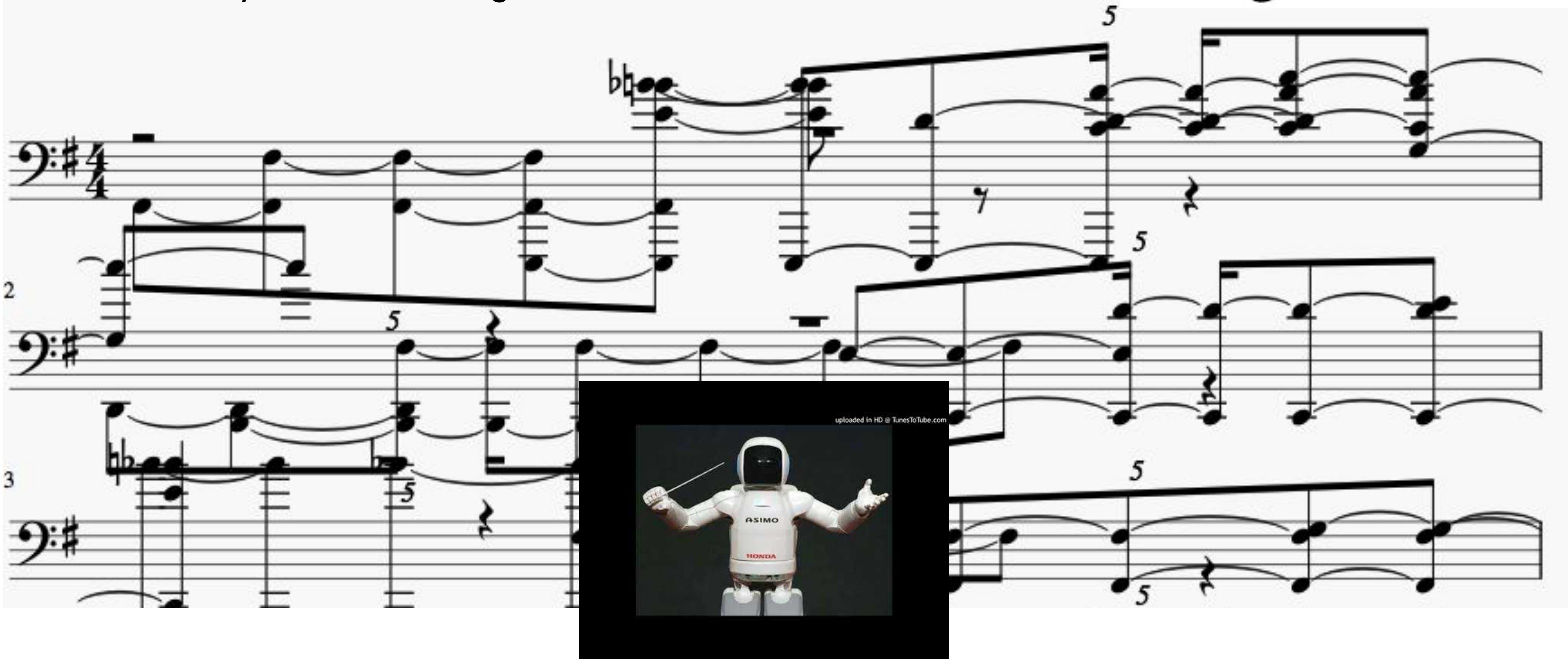
Generated output from training on dance music



Robot image courtesy of trendhunter.com.

Dance - Piano

Generated output from training on dance music



Piano

Robot image courtesy of trendhunter.com.

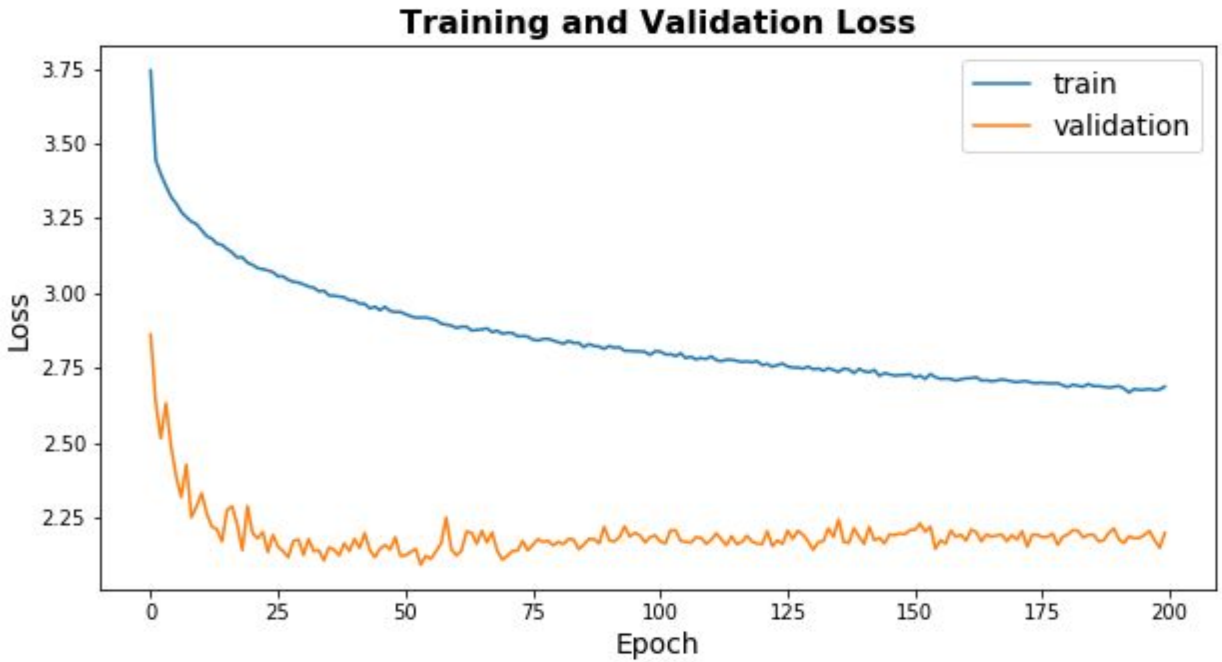
Datasets

Scraped by genre from various websites

Genre	# Midi	# Notes	# Unique Notes	Source
Celtic	338	159,789	78	Tadpole Tunes
Dance	200	309,967	663	MidiWorld
Jazz	15	9,326	292	MidiWorld
Game	91	51,177	358	Final Fantasy soundtracks*

**See blog post by Sigurour Skuli, Towards Data Science.*

Evaluation Loss



A Model With Memory

Recurrent (esp. LSTM) model an essential component of:

- Sound and speech recognition
- Time series prediction: traffic, recommender systems, stock movement
- Natural Language Processing (NLP): machine translation, chatbots
- Digital assistants

Recurrent Network

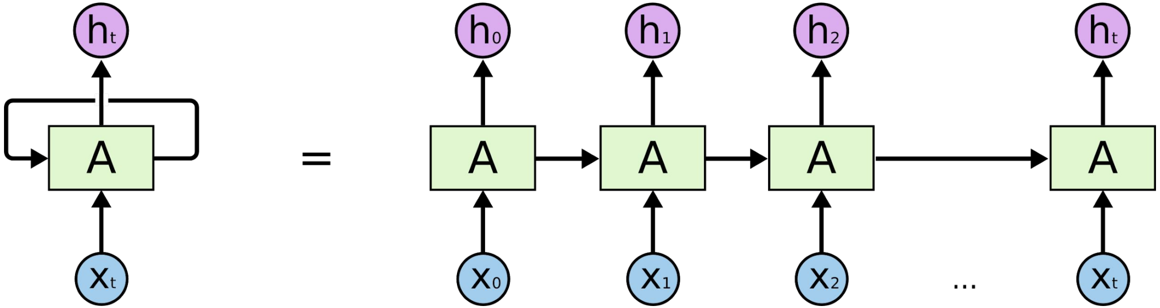


Diagram courtesy of "colah's blog" at https://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM Diagram (2)

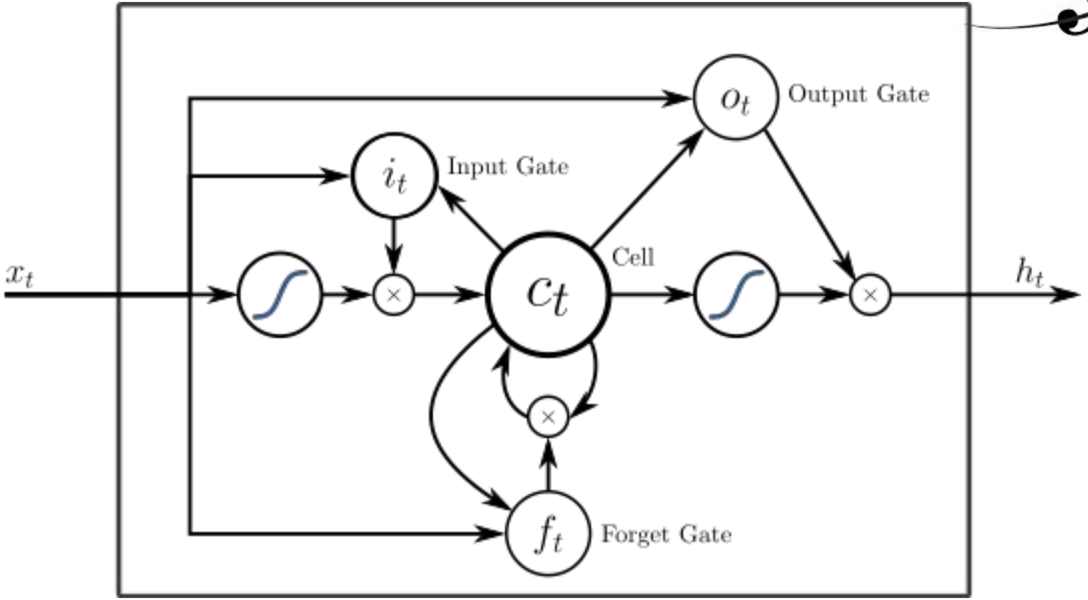


Diagram courtesy of "colah's blog" at [*https://colah.github.io/posts/2015-08-Understanding-LSTMs/*](https://colah.github.io/posts/2015-08-Understanding-LSTMs/)

Creativity in AI

A long disputed and contentious question: can AI be creative?

- “Remixing” precedent with a dose of stochasticity
- Potential to generate new thoughts and ideas unbounded by the human experience

Lessons Learned

- Successfully implemented a functional AI music generator
- Tested the audio and generative capabilities of neural networks
- Utilized various audio format preprocessing

Resources

Dorsey, Brannon. "Using Machine Learning to Create New Melodies." https://brangerbriz.com/. 10 May 2017.

Nayebi, Aran. "GRUV: Algorithmic Music Generation using Recurrent Neural Networks." Stanford University. 2015.

Skúli, Sigurður. "How to Generate Music using a LSTM Neural Network in Keras." www.towardsdatascience.com. December 7, 2017.

Brownlee, Jason. "Stacked LSTM Networks." https://machinelearningmastery.com. August 18, 2017.

Brownlee, Jason. "Understand the Difference Between Return Sequences and Return States for LSTMs in Keras." https://machinelearningmastery.com. October 24, 2017.

"Understanding LSTM Networks." Colah's Blog. https://colah.github.io. 27 August 2015.

Goodfellow, Ian. "Deep Learning." MIT Press. http://www.deeplearningbook.org/. 2016.

"Magenta." Tensorflow. Magenta.tensorflow.org.