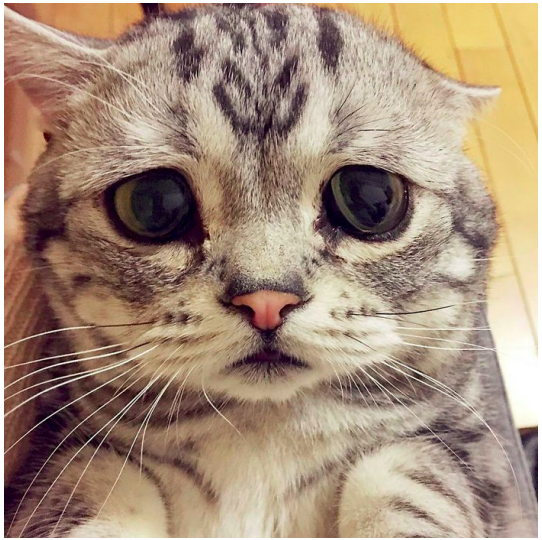
4. RNN/LSTM

Computational Music Creativity

Save the Pets




THEY'RE EATING THE CATS!
LIVE ABC NEWS • 2024 PRESIDENTIAL DEBATE •


THIS IS
SPA...SPRINGFIELD
imgflip.com

Divide into 4 teams

- Cat rescue
- Dog rescue
- Hamster rescue
- Guinea pig rescue

Save the Pets rules

- Valerio asks questions

Save the Pets rules

- Valerio asks questions
- Only one person answers

Save the Pets rules

- Valerio asks questions
- Only one person answers
- Others from your team can tag

Save the Pets rules

- Valerio asks questions
- Only one person answers
- Others from your team can tag
- If you answer well, you save a pet

Save the Pets rules

- Valerio asks questions
- Only one person answers
- Others from your team can tag
- If you answer well, you save a pet
- Sometimes, you'll have time to reflect with your team

Save the Pets rules

- Valerio asks questions
- Only one person answers
- Others from your team can tag
- If you answer well, you save a pet
- Sometimes, you'll have time to reflect with your team
    - Buzz in to answer

Save the Pets rules

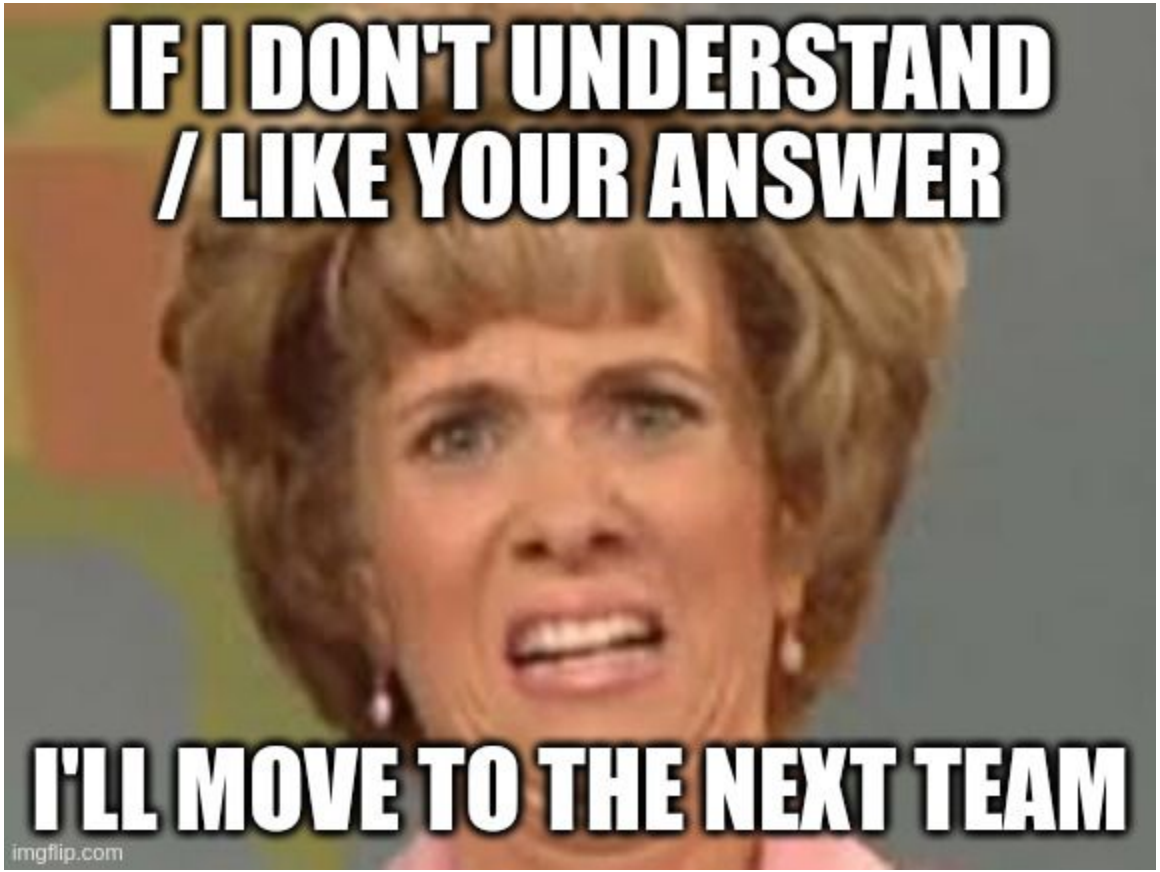
- Valerio asks questions
- Only one person answers
- Others from your team can tag
- If you answer well, you save a pet
- Sometimes, you'll have time to reflect with your team
    - Buzz in to answer
    - Right answer -> 2 pets get saved

Save the Pets rules

- Valerio asks questions
- Only one person answers
- Others from your team can tag
- If you answer well, you save a pet
- Sometimes, you'll have time to reflect with your team
    - Buzz in to answer
    - Right answer -> 2 pets get saved
    - Wrong answer -> 2 pet get killed

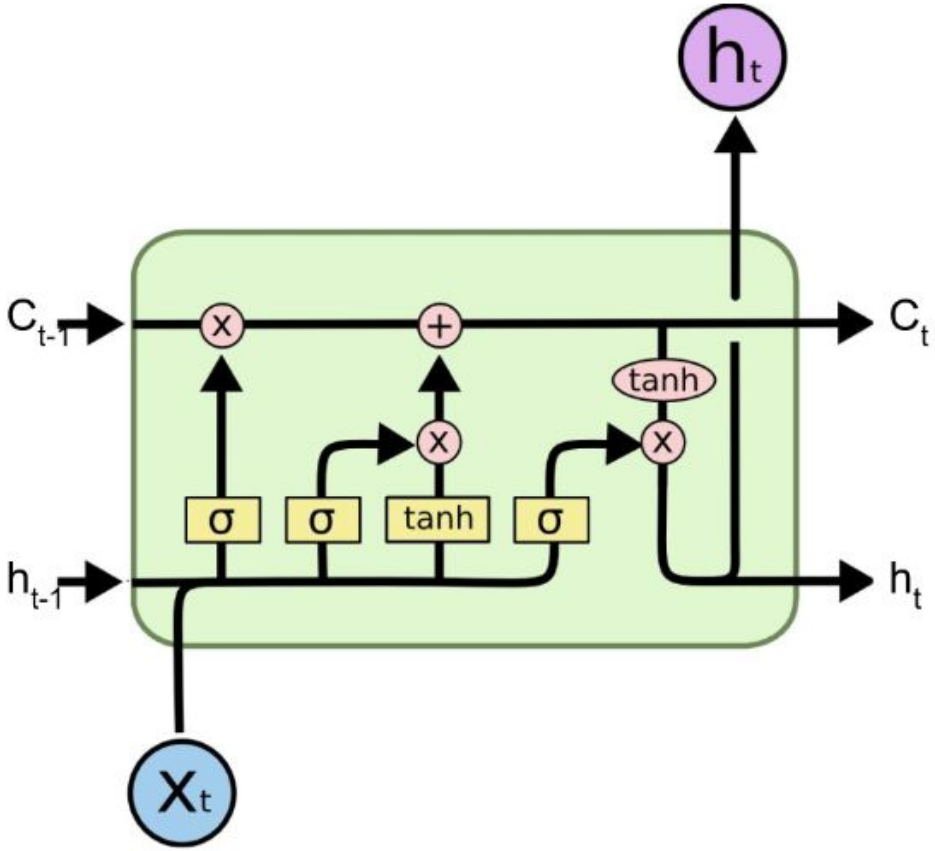
What I like

- Straight to the point
- No fluff
- Answer what I ask


IF I DON'T UNDERSTAND
/ LIKE YOUR ANSWER
I'LL MOVE TO THE NEXT TEAM
imgflip.com

Real-time scores


Pet Savers
Pearlyarts


h_t
C_{t-1}
C_t
tanh
σ
σ
tanh
σ
h_{t-1}
h_t
X_t

My experience with RNN/LSTMs

- Great for melody / chord generation

My experience with RNN/LSTMs

- Great for melody / chord generation
- Quality music data >> quantity of music data

My experience with RNN/LSTMs

- Great for melody / chord generation
- Quality music data >> quantity of music data
- Work well to generate a phrase. Beyond that -> random walk

My experience with RNN/LSTMs

- Great for melody / chord generation
- Quality music data >> quantity of music data
- Work well to generate a phrase. Beyond that -> random walk
- Handle simple poliphony well

Tips to use LSTMs

- Use dropout layers (LSTM overfits)

Tips to use LSTMs

- Use dropout layers (LSTM overfits)
- Craft music-aware representation
  (MIDI sucks)

Tips to use LSTMs

- Use dropout layers (LSTM overfits)
- Craft music-aware representation
  (MIDI sucks)
- Play with *temperature*

Tips to use LSTMs

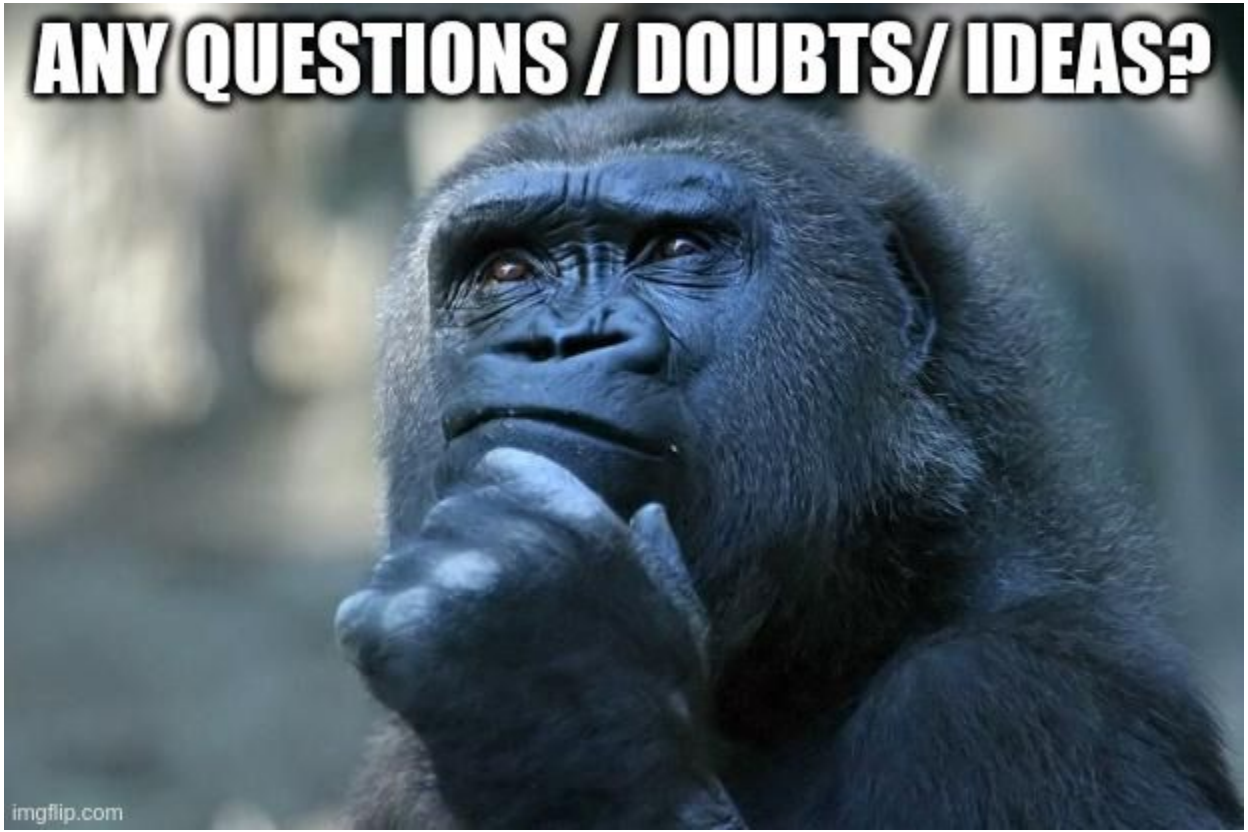
- Use dropout layers (LSTM overfits)
- Craft music-aware representation (MIDI sucks)
- Play with *temperature*
- Objective evaluation sucks -> use humans

Tips to use LSTMs

- Use dropout layers (LSTM overfits)
- Craft music-aware representation (MIDI sucks)
- Play with *temperature*
- Objective evaluation sucks -> use humans
- Genre-specific LSTM >> one LSTM to rule them all

Tips to handle data with LSTMs

- Transpose all your melodies to C / Amin
- Augment data transposing to all keys


ANY QUESTIONS / DOUBTS/ IDEAS?
imgflip.com

Activity 1: LSTM Variation

How do you use an LSTM melody generation model to evaluate melody inpainting variations?

Instructions:

- Work in groups (5 people)
- 10’ to come up with solution
- 10’ to discuss together

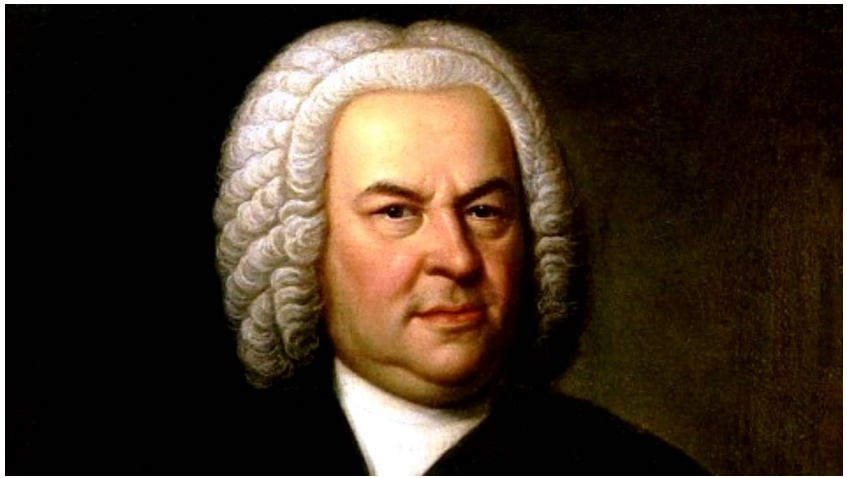
BachBot

BachBot goal

- Generate chorales in the style of Bach from scratch
- Harmonize melodic line (e.g., S), with other 3 voices (A, T, B) in the style of Bach

Corpus

- Train on 350+ Bach Chorales



Corpus pre-processing

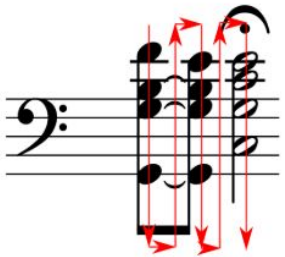
- Transpose (Cmaj / Amin)

Corpus pre-processing

- Transpose (Cmaj / Amin)
- Quantize

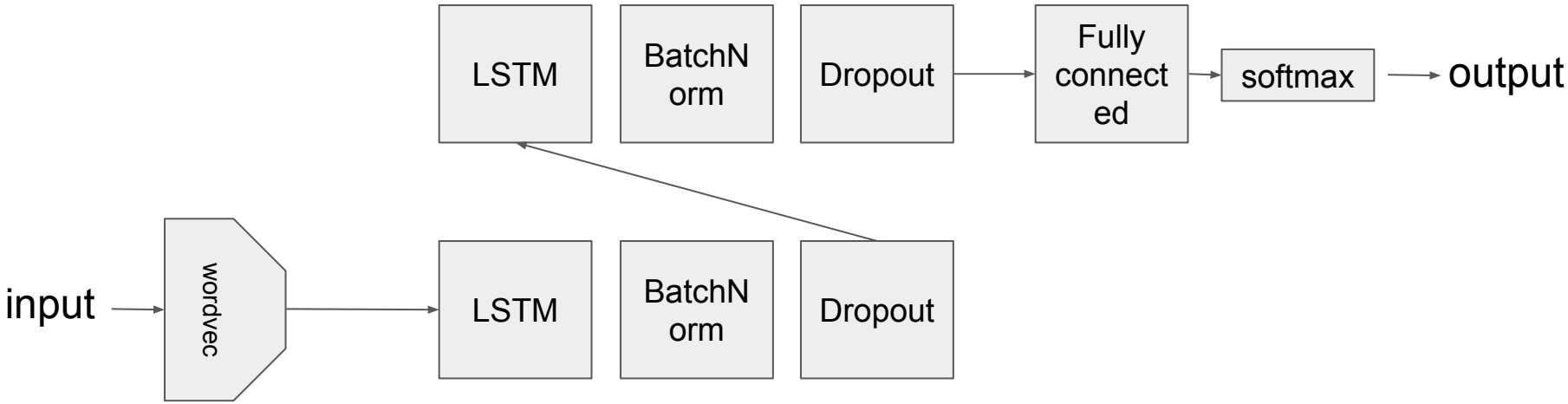
Corpus pre-processing

- Transpose (Cmaj / Amin)
- Quantize
- Tokenize individual notes in SATB, with delimiter (| | |)



```
START
(65, False)
(59, False)
(55, False)
(43, False)
|||
(64, False)
(59, True)
(55, True)
(43, True)
|||
(.)
(64, False)
(60, False)
(55, False)
(48, False)
|||
END
```

Model architecture



Model training

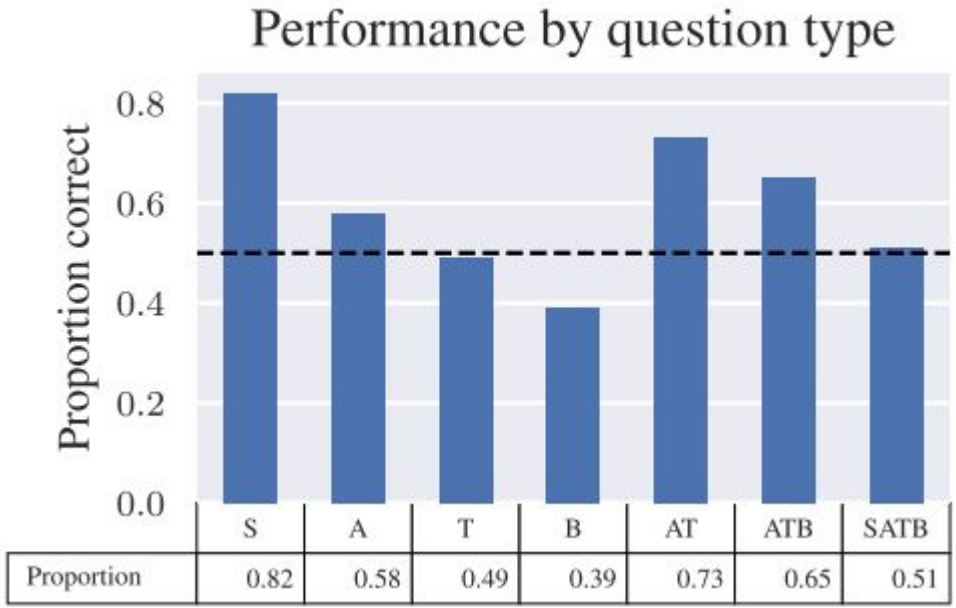
- Cross-entropy loss between predicted distribution and target distribution
- Teacher forcing

Hyperparameter fine tuning

- Grid search

Parameter	Values Searched
num_layers	$\{1, 2, 3, 4\}$
rnn_size	$\{128, 256, 384, 512\}$
wordvec	$\{16, 32, 64\}$
seq_length	$\{64, 128, 256\}$
dropout	$\{0.0, 0.1, 0.2, 0.3, 0.4, 0.5\}$

Evaluation



	S	A	T	B	AT	ATB	SATB
Proportion	0.82	0.58	0.49	0.39	0.73	0.65	0.51

Activity 2: Evaluation

Provide evaluation strategies to assess the quality of music generated from an LSTM.
For each:

- Pros and cons
- Limitations

Instructions:

- Work in groups (5 people)
- 10’ to come up with solution
- 10’ to discuss together

Assignment 3: Rock your LSTM

Implement an LSTM to automatically generate sequences of chords for rock music.

Deadline: 25 January at midnight