

Generowanie muzyki używając rekurencyjnych sieci neuronowych

Wojciech Buczek

9 czerwca 2019

Opis projektu

Celem programu jest komponowanie utworów poprzez uczenie się na podstawie utworów już istniejących. Na wejściu program dostaje pliki MIDI do celów uczenia się, jego zadaniem jest stworzenie odpowiedniego pliku MIDI z ładną muzyką.

Metody sztucznej inteligencji

Do rozwiązania problemu zostały użyte rekurencyjne sieci neuronowe Long short-term memory (LSTM). W muzyce bardzo istotna jest kolejność nut, więc rekurencyjna sieć jest dobrym rozwiązaniem. Sieci LSTM rozwiązują 'Vanishing gradient problem', czyli sytuacje, gdzie gradient, który jest back-propagowany może zniknąć (tzn. dążyć do 0) lub 'eksplodować' (czyli dążyć do nieskończoności).

Architektura sieci neuronowej

- LSTM
- LSTM
- Dropout(0.2)
- Dense

Funkcja aktywacji: softmax

Funkcja straty: Categorical crossentropy

Optimizer: RMSprop

Wkład własny

- Zebranie danych
- Przekształcenie danych dla wejścia do sieci neuronowej
- Przygotowanie odpowiedniej sieci neuronowej, dobranie architektury, wejścia
- Uczenie sieci
- Generowanie muzyki używając modelu
- Zapisanie dźwięków w odpowiednim formacie, np. midi lub mp3.

Narzędzia

- Python 3
- Tensorflow 2.0.0-alpha (z obsługą CUDA)
- Keras
- Python-midi
- Tqdm

Wyniki

Po odpowiednio długim przetrenowaniu sieci otrzymaliśmy ładne, przyjemne dla ucha melodie. Zapisane są w folderze 'generated', ale nic nie przeszkadza, aby wygenerować nowe.

Parametry oraz architektura sieci neuronowej zostały tak dobrane, aby utwór brzmiał jak najładniej oraz jak najlepiej malał loss. Przy obecnych ustawieniach wydaje się spełniać oba warunki.