

Zadanie 1

wtorek, 8 marca 2022

11:30

ARTIST SIMILARITY WITH GRAPH NEURAL NETWORKS

Jest to praca naukowa opisująca sieć GNN służącą do wyszukiwania podobieństw między artystami. Wykorzystywany jest w systemach rekomendacji. Podobieństwo możemy badać na podstawie harmonii, melodii czy też współwystępowanie pozycji na listach odtwarzania. Wykorzystują metody grafowe oraz embedding. Funkcja loss - triplet loss ma na celu przybliżyć podobnych artystów, a oddalić negatywne przypadki. Wykorzystywany jest dataset OLGA. Porównując wyniki dla GNN oraz DNN, grafy lepiej prognozują prawdopodobieństwo artystów.

Leveraging Hierarchical Structures for Few-Shot Musical Instrument Recognition (ISMIR 2021)

Rozpoznawanie instrumentów muzycznych. Dataset jest ograniczony, jednak chcielibyśmy nauczyć naszą sieć rozpoznawać nowy dźwięk jedynie za pomocą kilku próbek. Wykorzystywany jest system hierarchii, który po wyznaczeniu hierarchii instrumentów umożliwi naukę systemu za pomocą kilku próbek nowej klasy który ma wspólnego przodka w naszej wyznaczonej hierarchii. Wykorzystujemy sieci prototypowe. Początkowo tworzymy sieć na najbardziej szczegółowym poziomie. Następnie wykorzystujemy predefiniowaną sieć do stworzenia nowego prototypu dla grubszych poziomów. Powtarzamy ten proces, co pozwala na rozpoznanie instrumentów na różnych poziomach. Obliczamy embedding vector dla każdej próbki z podzbioru zbioru S zawierającego wszystkie przykłady dla danej klasy k . Tworzymy rozkład prawdopodobieństwa na zbiorze klas dla danego zapytania. Struktura ma postać drzewa. Hierarchiczne sieci prototypowe są rozszerzeniem sieci prototypowych o hierarchiczne uczenie wielozadaniowe (etykietowanie zapytań).

BebopNet

Generowanie spersonalizowanej muzyki. Generowanie muzyki porównywane jest do NLP. Na podstawie znaków (nut) przewidywany jest kolejny znak (nuta). Aby stworzyć spersonalizowane improwizacje jazzowe przeprowadzane jest badania oraz uczenie preferencji użytkownika. a następnie zoptymalizowanie generowania

...nauczenie preferencji użytkownika, generowanie muzyki. Aby nauczyć się preferencji użytkownika wykorzystywane jest recurrent regression model. Nuta reprezentowana jest przez 2 wartości (pitch-duration). BebopNet trenujemy za pomocą datasetu solo jazzowych z próbkami w formacie musicXML. Dostarczane dane to: pitch, duration, offset i chord. Model może zostać różnie zaimplementowany: RNN, CNN, LSTM... Dane reprezentujemy przez warstwy embeddingu. Preferencje użytkowników pobierane są na podstawie datasetu - 124 improwizacji (oryginalne jak i wygenerowane przez BebopNet). Wykorzystując transfer learning, tworzymy model preferencji użytkownika składający się z tych samych warstw jak BebopNet, bez końcowych warstwy fully-connected. Wagi warstw embeddingu pozostają zamrożone.

Mingus

Generowanie muzyki (jazz).

Podczas generowania muzyki możemy wykorzystać techniki wykorzystywane w NLP MINGUS przyjmuje na wejściu odpowiednia sekwencje z wartościami takimi jak Pitch, Duration, Chord, Next Chord, Bass, Beat, Offset

Zbudowany jest z 2 równoległych modeli, do przewidzenia wysokości dźwięku oraz czasu trwania.

Modele te są takie same, różnią się jedynie wyjściem. Oba modele wykorzystują funkcje crossentropy oraz SGD

Generowanie muzyki polega na dostarczeniu sekwencji nut o zadanej długości.

Następnie wejście dzielone jest na sekwencje wysokości dźwięku oraz czasu trwania a następnie wprowadzone do odpowiedniego modelu. Wynikiem działania jest token o największym prawdopodobieństwie z wyjścia modelu. Następnie jest on dołączany do naszej wygenerowanej sekwencji

