

UNIwersytet IM. ADAMA MICKIEWICZA W POZNANIU
WYDZIAŁ MATEMATYKI I INFORMATYKI

Cezary Adam Pukownik

Kierunek: Analiza i przetwarzanie danych

Specjalność: Uczenie maszynowe

Numer albumu: 444337

**Generowanie muzyki
przy pomocy głębokiego uczenia**

Music generation with deep learning

Praca licencjacka
napisana pod kierunkiem
dr hab. Tomasza Góreckiego

POZNAŃ 2020

Poznań, dnia

OŚWIADCZENIE

Ja, niżej podpisany Cezary Pukownik, student Wydziału Matematyki i Informatyki Uniwersytetu im. Adama Mickiewicza w Poznaniu oświadczam, że przedkładaną pracę dyplomową pt: "Generowanie muzyki przy pomocy głębokiego uczenia", napisałem samodzielnie. Oznacza to, że przy pisaniu pracy, poza niezbędnymi konsultacjami, nie korzystałem z pomocy innych osób, a w szczególności nie zlecałem opracowania rozprawy lub jej części innym osobom, ani nie odpisywałem tej rozprawy lub jej części od innych osób.

Oświadczam również, że egzemplarz pracy dyplomowej w wersji drukowanej jest całkowicie zgodny z egzemplarzem pracy dyplomowej w wersji elektronicznej.

Jednocześnie przyjmuję do wiadomości, że przypisanie sobie, w pracy dyplomowej, autorstwa istotnego fragmentu lub innych elementów cudzego utworu lub ustalenia naukowego stanowi podstawę stwierdzenia nieważności postępowania w sprawie nadania tytułu zawodowego.

[TAK]* - wyrażam zgodę na udostępnianie mojej pracy w czytelni Archiwum UAM

[TAK]* - wyrażam zgodę na udostępnianie mojej pracy w zakresie koniecznym do ochrony mojego prawa do autorstwa lub praw osób trzecich

*Należy wpisać TAK w przypadku wyrażenia zgody na udostępnianie pracy w czytelni Archiwum UAM, NIE w przypadku braku zgody. Niewypełnienie pola oznacza brak zgody na udostępnianie pracy.

.....

Spis treści

Streszczenie	7
Abstract	9
Rozdział 1. Wstęp	11
1.1. Zastosowania uczenia maszynowego w muzyce	11
1.2. Modele generatywne	11
1.3. Muzyka symboliczna, a muzyka	11
1.4. Cele tej pracy.	12
Rozdział 2. Reprezentacja muzyki	13
2.1. Podstawowe koncepcje	13
2.1.1. Nuta	13
2.1.2. Skala	13
2.1.3. Akord	13
2.1.4. Utwór	14
2.2. Reprezentacja muzyki	14
2.2.1. Zapis klasyczny - pięciolinia	14
2.2.2. Tabulatura	14
2.2.3. Pianoroll	14
2.2.4. Tekstowa	15
Rozdział 3. Sieci neuronowe	17
3.0.1. Wstęp do sieci neuronowych, definicje wzory itp.	17
3.0.2. Autoencodery	17
3.0.3. LSTM	17
3.0.4. GAN	18
Budowanie generatora muzyki	19
3.0.5. Przygotowanie danych	19
3.0.6. Architektura sieci neuronowej	19
3.0.7. Proces treningowy	19
3.0.8. Przykłady wygenerowanej muzyki	20
Podsumowanie	21
Bibliografia	23

Streszczenie

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut fermentum libero libero. Duis a magna arcu. Nam sit amet porta odio. Cras sit amet euismod elit. Etiam a turpis eget magna pharetra malesuada. Vivamus accumsan leo eget turpis efficitur, non interdum tortor pretium. Maecenas at massa nec elit imperdiet sagittis. Maecenas pellentesque libero et risus aliquam consectetur.

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut fermentum libero libero. Duis a magna arcu. Nam sit amet porta odio. Cras sit amet euismod elit. Etiam a turpis eget magna pharetra malesuada. Vivamus accumsan leo eget turpis efficitur, non interdum tortor pretium. Maecenas at massa nec elit imperdiet sagittis. Maecenas pellentesque libero et risus aliquam consectetur.

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

Wstęp

Uczenie maszynowe w ostatnich latach mocno zyskało na popularności. Zastosowania i możliwości różnych algorytmów Machine Learning czasami przekraczają nasze wyobrażenie o tym, co komputer może zrobić. Niektóre aplikacje potrafią wręcz zaskoczyć użytkowników tym, co potrafią zrobić. Wśród takich aplikacji znajdują się takie, które potrafią przewidywać następne wartości akcji giełdowych, rozpoznawać na filmie obiekty w czasie rzeczywistym czy nawet prowadzić samochód. Algorytmy wyuczone proponują nam spersonalizowane reklamy, czy produkty na podstawie naszych upodobań. Najczęstsze zastosowania dotyczą przetwarzania obrazów lub tekstu, natomiast zastosowania w przetwarzaniu muzyki są niszowe i rzadko spotykane.

1.1. Zastosowania uczenia maszynowego w muzyce

Wśród najbardziej rozwiniętych zastosowań uczenia maszynowego w muzyce, można wymienić algorytmy polecenia utworów w portalach streamingowych takich jak Spotify czy Tidal. Algorytmy potrafią znajdować podobne do siebie utwory i polecać je nam na podstawie naszych preferencji.

1.2. Modele generatywne

Jednym z najnowszych modeli sieci neuronowych są sieci generatywne. Wychodzą one poza standardowe zastosowania klasyfikacji i regresji. Modele generatywne, uczą się ze zbioru uczącego najważniejszych ale i ogólnych cech i potrafią reprodukcować podobne wyniki.

1.3. Muzyka symboliczna, a muzyka

Należy rozróżnić dwa pojęcia, które są od siebie różne w podstawowych założeniach. Muzyka symboliczna, i muzyka odegrana. Muzyka symboliczna, jest to utwór zapisany, skomponowany ale na papierze. Przedstawia to muzykę jako koncepcję, taki przepis na utwór. Taka muzyka zapisywana jest klasycznie

na pięciolinii, czy komputerowo przy pomocy protokołu MIDI. Druga muzyka, jest to muzyka już odebrana, która nie przechowuje informacji o tym jak zagrać, czy odtworzyć utwór muzyczny, ale brzmienie tego utworu jako fala dźwiękowa. W tej pracy będę opisywał przede wszystkim generowanie muzyki symbolicznej.

1.4. Cele tej pracy.

Celem tej pracy, jest zastosowanie technik głębokiego uczenia, do generowania muzyki. Jest to bardzo ogólny cel, ponieważ muzyką może być prosta melodia oparta na kilku dźwiękach grana przez jeden instrument ale również aranżacja orkiestralna na wiele instrumentów, które razem współgrają i wybrzmiewają jako jeden pełny utwór.

Reprezentacja muzyki

Muzyka jesy

2.1. Podstawowe koncepcje

2.1.1. Nuta

Każdy utwór muzyczny składa się nut. Nuta jest podstawowym obiektem w muzycznym słowniku. Każda nuta ma dwa parametry, wartości oraz wysokości. Wartość nuty określa jak długo będzie ona trwać w czasie, relatywnie do pozostałych nut. Wysokość noty oznacza z jaką częstotliwością fala dźwiękowa tej nuty ma wybrzmieć. Częstotliwości te są nazwane literami alfabetu ABCDEFG lub w zapisie polskim AHCDEFG.

2.1.2. Skala

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut fermentum lorem libero. Duis a magna arcu. Nam sit amet porta odio. Cras sit amet euismod elit. Etiam a turpis eget magna pharetra malesuada. Vivamus accumsan leo eget turpis efficitur, non interdum tortor pretium. Maecenas at massa nec elit imperdiet sagittis. Maecenas pellentesque libero et risus aliquam consectetur.

2.1.3. Akord

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

2.1.4. Utwór

Aenean malesuada interdum hendrerit. Integer quis nisl et neque iaculis dapibus at in metus. Cras pretium bibendum magna at aliquet. Integer aliquet cursus augue, efficitur sollicitudin felis fringilla efficitur. Vivamus euismod bibendum justo, vitae suscipit nunc mattis a. Sed egestas porttitor velit, sit amet volutpat tortor suscipit vitae. Nulla nec dignissim mauris. Curabitur maximus viverra mollis. Suspendisse molestie turpis sit amet turpis interdum viverra ac eu lorem. Suspendisse iaculis ultricies ante, a condimentum odio congue nec. Integer varius lobortis diam, eget scelerisque nisl mattis at.

2.2. Reprezentacja muzyki

2.2.1. Zapis klasyczny - pięciolinia

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

2.2.2. Tabulatura

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

2.2.3. Pianoroll

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum

eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

2.2.4. Tekstowa

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

Sieci neuronowe

Tutaj będzie opisane, dlaczego sieci neuronowe, radzą sobie lepiej w produkcji muzyki niż inne modele. Oraz jakie modele są odpowiednie do pewnych zastosowań, JAZZ - LSTM, bardziej ustrukturyzowana - VAE itp.

3.0.1. Wstęp do sieci neuronowych, definicje wzory itp.

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

3.0.2. Autoencodery

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

3.0.3. LSTM

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et

malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

3.0.4. GAN

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

Budowanie generatora muzyki

W tym rozdziale opiszę w jaki sposób zbudowałem swój własny generator muzyki, jak przechodził proces uczenia, jakie próbki udało mi się wygenerować. Opis kodu który napisałem.

3.0.5. Przygotowanie danych

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

3.0.6. Architektura sieci neuronowej

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

3.0.7. Proces treningowy

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et

malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

3.0.8. Przykłady wygenerowanej muzyki

Proin ac dui orci. Cras nec elit eleifend lacus eleifend gravida. Ut placerat lacinia dolor non viverra. Curabitur rhoncus sit amet nibh sed malesuada. Integer iaculis eros venenatis, tempor enim non, sollicitudin sapien. Vestibulum eu scelerisque erat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla at felis massa. Ut est arcu, rhoncus ac tincidunt vel, consequat eu sem. Aliquam neque orci, lacinia molestie enim fermentum, ullamcorper congue mauris. Phasellus pellentesque, ante nec ultricies porta, erat erat placerat ante, vitae vehicula ipsum enim id ante. Donec malesuada tortor id ornare mattis. Nulla nec augue at augue dictum aliquet.

Podsumowanie

Ostateczne wnioski, czy muzyka generowana komputerowa da się lubić? Czy to pozytywnie wpłynie na przemysł muzyczny? Tak i nie. Może złużyć jako inspiracja dla muzyków, proces wspierający. Z drugiej strony może obniżyć koszty produkowania muzyki pop, która i tak jest już bardzo powtarzalna. Czy sieci neuronowe nauczą się produkować Hity?

Bibliografia

- [1] Briot, J.P., Hadjeres, G., Pachet, F.D. (2019): *Deep Learning Techniques for Music Generation - A Survey*. *arXiv:1709.01620v3*
- [2] Goodfellow, I., Bengio, Y., Courville, A. (2016): *Deep Learning*. MIT Press.
- [3] Zocca, V., Spacagna, G., Slater, D., Roelants, P. (2018): *Deep Learning. Uczenie głębokie z językiem Python*. Helion.