



Generisanje Lofi muzike uz pomoć neuronskih mreža (RNN i CNN)

Vukan Antić

225/2018

Petar Nikić

049/2018

Matematički fakultet
Univerzitet u Beogradu

1. Uvod

Šta je Lofi?

- Low-fidelity music
- Sastoji se od 3 ključna dela:
 - Muzickog uzorka (Sample)
 - Buhnjeva
 - Pozadinskih zvukova (npr. padanje kiše)
- **Ideja** : Generisanje muzičkog uzoraka korišćenjem mreže, i onda se dodaju buhnjevi koji su sinhronizovani sa njim

Obrada podataka

- Za obradu uzorka korišćena je biblioteka ***Music21***
- Svaka pesma se sastoji iz niza nota i akorda(Note, Chord)
- Uz pomoć biblioteke, razlaže se pesma na elemente, i formira se jedan veliki niz nota
 - ❑ Pesme su slične prirode, samim tim ne pravi veliku razliku da li će se sve note složiti u jedan niz, ili detaljno odvajati po pesmama
 - ❑ Postojao je pokušaj da se razdvoje različite pesme unutar istog niza korišćenjem padding-a, ali nije pravilo nikakvu razliku

Rest

- Pauza (Rest) pored note i akorda, predstavlja ključni deo svake pesme, ali je problematičan za obradu
 - Zbog veličine uzorka koji je korišćen za treniranje modela i dužine samih pesama, pauze čine većinu od svih mogućih izbora
 - Problem nastaje onda, jer se model previše navikne na pauze, i onda dobijamo rezultate gde pauza čini većinu pesme, ili je cela pesma

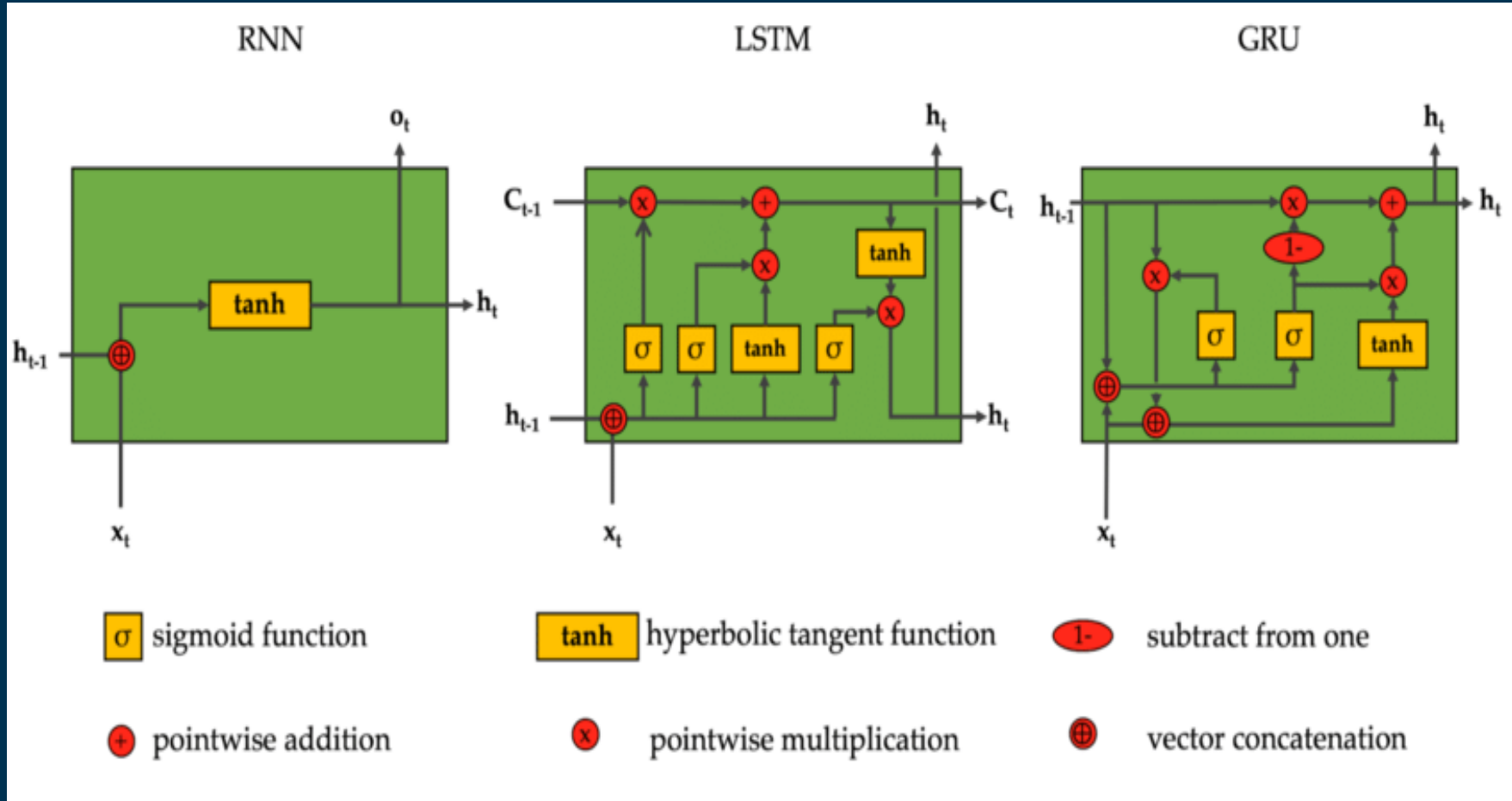
Priprema za treniranje

- Korišćenjem niza svih noti, organizujemo podatke za trening
- Niz nota dužine `sequence_length` nam predstavlja jednu instancu, a predviđenu vrednost nam predstavlja nota koja se u nizu svih nota nalazi posle njega
- Primer :
 - ❖ Ako niz svih nota ide `[A, B, C, C, D, E,]`, i `sequence_length = 5`
 - ❖ `X1 = [A, B, C, C, D]` i `Y1 = E`
- Što je veći `sequence_length`, to model više uči dugotrajnu strukturu (eksperimentalnim metodama određen)

2. Modeli

Rekurentne neuronske mreže

- Pored trenutne ulazne vrednosti, rekurentne mreže uzimaju u obzir i rezultate prethodnih ulaza
- Ovakve mreže su savršene za generisanje muzike, jer pri izboru sledeće note, ne želimo da uzimamo samo prethodnu notu, već prethodnih n nota u obzir
- Vrste rekurentnih mreža:
 - I. Prost RNN
 - II. GRU
 - III. LSTM



RNN

- Prost RNN predstavlja klasični oblik RNN-a, ali se ne koristi zbog:
 - U toku backpropagation-a nastaje problem *nestajućeg gradijenta*
 - Posle nekog vremena gradijent postane toliko mali, da više ne utiče na učenje
- LSTM i GRU rešavaju ovaj problem uvođenjem kapija (*gate*), koje predstavljaju kontrolisano menjanje stanja
- LSTM funkcioniše koristeći koncept *ćelije*, na osnovu koje se određuje vrednost skrivenog sloja, i 3 kapije koje kontrolišu promene (forget, input, output)
- GRU koristi 2 kapije (update i reset), i ne koristi koncept *ćelije*, zbog toga je jednostavniji, ali ima slične karakteristike kao LSTM

Konvolutivne neuronske mreže

- Konvolutivne neuronske mreže karakteriše naizmenični niz 1D *konvolucija*, i 1D *max agregacija*
- Ovakve mreže formiraju narednu vrednost uzimajući u obzir prethodne vrednosti sekvence, i zato one isto predstavljaju dobro rešenje našeg problema

Generisanje pesme

- Da bi se generisala nova pesma, najpre se napravi nasumičan niz noti dužine `sequence_length`
- Uz pomoć tog niza i modela, generišemo narednu notu
- Odbacujemo prvu notu sa nasumičnog niza, i dodajemo generisanu notu na kraj istog
- Generišemo pesmu dužine koja je unapred zadata (500 noti)

3. Rezultati

Neki od rezultata

GRU



LSTM



CNN



Naši favoriti (Odsečci pesama)



Reference

1. <https://towardsdatascience.com/how-to-generate-music-using-a-lstm-neural-network-in-keras-68786834d4c5>
2. <https://machinelearningmastery.com/text-generation-lstm-recurrent-neural-networks-python-keras/>
3. <https://ai.plainenglish.io/building-a-lo-fi-hip-hop-generator-e24a005d0144>
4. <https://www.analyticsvidhya.com/blog/2020/01/how-to-perform-automatic-music-generation/>
5. Alex Graves. „Generating Sequences With Recurrent Neural Networks”.
<https://arxiv.org/abs/1308.0850>
6. DeepMind „WaveNet: A Generative Model for Raw Audio“
<https://arxiv.org/abs/1609.03499>



Hvala na pažnji.