**Report**

Assignment 3, Deep Learning

by Carmel Shablin 305812661

and Gregory Koushnir 321889479

## Task

Implement a recurrent neural net to generate lyrics from one word based on the provided melody.

## Dataset

The dataset includes a .zip file containing all the MIDI files of the participating songs and two .csv files with all the lyrics of the of the participating songs, preliminary divided to 600 train and 5 test samples.

Most of the songs’ lyrics are in English and of different music genres including but not limited to POP, RAP and Rock. Lyrics are lowercased, splitted in rows and contain a title and name of the artist performing the song.

For the purpose of training of the recurrent neural network, the train set was randomly divided to train and validation sets in 0.9 and 0.1 ratio respectively.

## Environment

Since none of the students has or has an access to GPU physically, the assignment was done in the Google Colab environment in order to speed up the training process. All of the shown results are taken from the Jupyter notebook.

## Method

The general method is to build a LSTM network based on the embeddings of the lyrics and concatenate it with selected features of the melody from the MIDI file.

### Preprocessing:

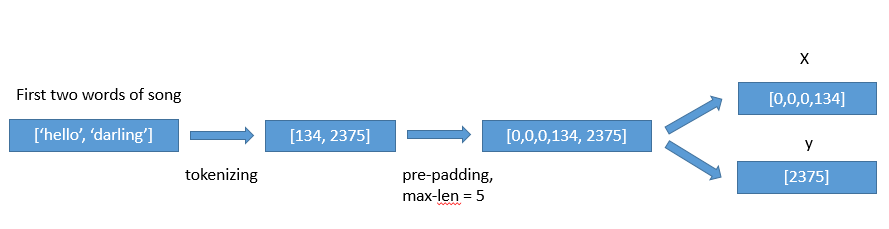
#### Lyrics

Text inside the parentheses, punctuation marks (with the exception of a single quote, for words like “I’ll, I’d, ‘cause” and alike) and the word “chorus” are removed. The “eof” token is added to the end of every lyrics to determine the end of the song.

Afterwards, sequences of word indexes as returned by the tokenizer are generated, starting from the first word and adding one by one all of the consecutive words of a song. Each sequence is pre-padded if needed with zeros and truncated to the given maximal length, resulting a n-gram sequence. Then, the n-gram is splitted to X and y, where X is all of the words in a n-gram sequence excluding the last one and y is last word of the sequence.

To build an embedding matrix, we use [GloVe](https://github.com/stanfordnlp/GloVe) vectors for word representation generated from Wikipedia 2014 + Gigaword 5 corpora including 6B tokens, 400K vocab. The vectors have a dimension of 300. Embedding matrix generation from the GloVe vectors for the lyrics corpora is done by utilizing the [glove\_embedding\_create](https://gist.github.com/shubham0204/531f75c37c49362c9b61a11865496177) script by shubhham0204, the code was changes so that in edition to the weight matrix generation, it will return the non\_words\_list, telling which words are not in the embedding matrix, with are removed from the lyrics before the model fitting. Practically, the remove\_words function was not used, because it was considered as incorrect to skip a word and link between non consequent words in the original lyrics and a lack of many slangs in the GloVe.

An example:



#### Melody

We tried tree python libraries: [MIDIFile](https://pypi.org/project/MIDIFile/), [pretty\_midi](https://nbviewer.jupyter.org/github/craffel/pretty-midi/blob/master/Tutorial.ipynb) and [mido](https://mido.readthedocs.io/en/latest/). The meta-data features of the midi files were extracted by the mido library and they consist of: key(A, Ab, B, Bb, C, C#,…), mode (major/minor), tempo and time signature data (numerator, denominator, clocks per tick and notated 32nd notes per beat). If a feature is not found by the parser, it is substituted to “-1” or “unknown” for numerical and categorical features respectively. The instruments were taken from midi tracks using the pretty\_midi library, the tracks count was summarized by the instrument type, from one of the following [categories](http://www.music.mcgill.ca/~ich/classes/mumt306/StandardMIDIfileformat.html#BMA1_4) {Drums, Piano, Chromatic Percussion, Organ, Guitar, Bass, Strings, Ensemble, Brass, Reed, Pipe, Synth Lead, Synth Pad, Synth Effects, Ethnic, Percussive, Sound Effects}.

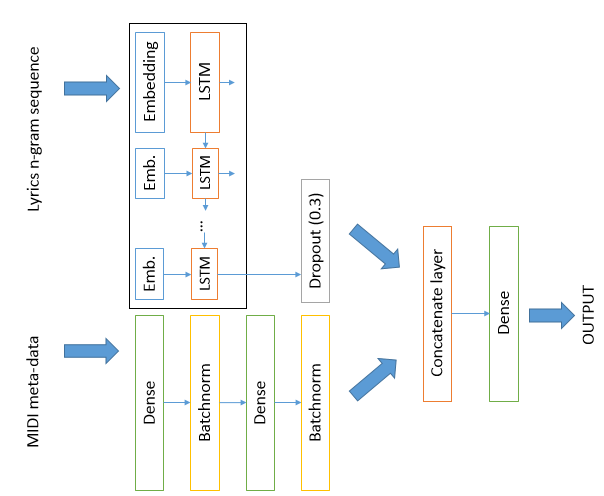
It was checked that the instruments features has at least 3 unique values, meaning that among the MIDIs all of the instrument categories are represented. But the performance of the model and the generated lyrics was dropped drastically when the instruments were added to the training. Thus, it was decided to remove the instruments from the samples data.

All of the numerical features were fit\_transformed with MinMax scaler on the train set and transformed on validation and test sets. The same action was done for the categorical features, but using the OneHotEncoder.

### Architecture

In this assignment, we were asked to implement two different architectures. Our choice is to compare early and late fusion architectures, where in early fusion the MIDI-metadata and the lyrics n-gram sequences are concatenated before running LSTM, and in the late fusion they are concatenated after the LSTM on lyrics and feed forward network on the MIDI meta-data. The same music representation features are used with sequences of a specific song.

The following architecture of the network is used for the Late fusion:



On each step, the network tries to predict the next word.

Lyrics n-gram sequence is an input for the trainable embedding layer, which has a weight of the previously generated embedding matrix as described earlier. The number of LST units is chosen to be as a number of the embedding dimension, which is 300 in our case. After the LSTM there is a dropout layer with 0.3 ratio. From the other input, the model gets melody features which go through the feed forward network with two dense layers (50 and then 10 units), ReLU activated and batch normalization after each dense layer. Lyrics and melody’s layers are then concatenated and connected to a final dense layer of the vocabulary size with the SoftMax activation, which gives us a representation of the distribution of prediction of the next word.

The model is trained with a parameter of maximal sequence length of 8 words, other values were checked manually, and this value was chosen based on the minimal validation loss results. Teacher forcing technique is used, so that every sequence is corrected with a ground truth lyrics’ word instead of the predicted one.

In total, there are 5,377,376 trainable weights.

The fit was done with the following parameters: batch size = 256, number of epochs = 5, we see that afterwards the validation loss stops to improve. Three callback functions were used:

* Early stopping for validation loss to interrupt training if there are no improvement for 3 epochs
* Model checkpoint, to save the weights of model with lowest validation loss
* TensorBoard callback for model performance logs dashboard

The loss that we are interested in is categorical-crossentropy. The optimizer that was used is Adam with the default learning rate of 1e-3, beta\_1=0.9, beta\_2=0.999, epsilon=1e-07. In addition to the loss metric, we are interested in accuracy.

After the training and finding the best model parameters, the model with the best validation loss is loaded for the evaluation on a test set.

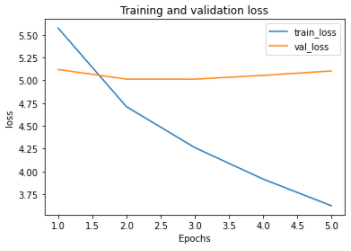
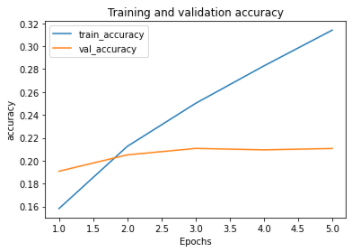
In the generation phase, given the first (or more) word of the lyrics and a MIDI features, the model predicts next word for the sequence which is picked from the vocabulary by the distribution of the output layer of the model. The sequence enrichment continues until the “eof” token is met.

One of the lexicological tweak that was checked, and declined afterward to not to be considered as “cheating”, is the language richness. We put as a constraint that the generated word has to be re-picked if it was met in the last 5 words sequence. The results were ambiguous. On the one hand, the generated lyrics seems to be sophisticated, but on the other hand there are missing expressions like “oh oh” and alike, which are used widely in nowadays songs. Thus, it was decided to exclude this tweak from the lyrics generation process.

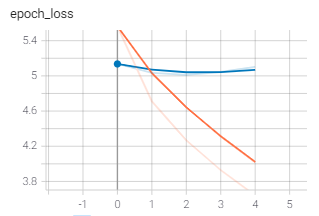
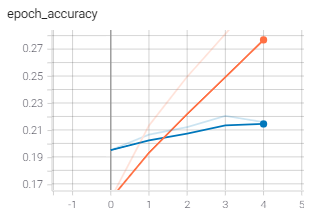
## Results

With the described architecture and method, the best epoch accuracy on a validation set is 5.0538 and accuracy of 0.2107.

Graphs that describe the accuracy and loss through the training process.

TensorBoard graphs (blue color represent the validation and orange is training):



The training takes ~60 second for the first epoch and ~27 seconds per consequent epochs. In total, it took 167 seconds to train. The measured test set accuracy is **0.25** and loss of 4.625.

Analyzing the song results:

* It seems that there is a bias of “love songs”, as trough all of the lyrics love motives pop-up
* It can be seen from a lyrics #3 of all the small things by blink 182, that the model stays consistent with language – once some Spanish words are generated, there is a great chance to see more of them
* There is no correlation for the lyrics nor line length, we got a distribution of lyrics and lines of different length for each input
* All in all we are very pleased with the results, as ,mostly (sometimes there are rows like “heaven don't can tell me”), they are linguistically permissible and sometimes even make sense

## Appendix A - Lyrics generation by late fusion:

|  |  |  |
| --- | --- | --- |
| Song: eternal flame by the bangles, Starting word: “close” | | |
| Generated lyrics #1 | Generated lyrics #2 | Generated lyrics #3 |
| [total number of words 329]  close in the face i'm lady out of gold  at all the morning light  i feel like phil i seen this girl  your own disappear  i make you see with your plays black side  and can't you stay trying to you  looking around  but i can't love to show you wait  my arms backwards  take your own  why say it's all right  from the good times  you let me gently  if you want me  let the night  keep in his kiss we can make you to  so you'll just get you  we'll feel alright  i know a way  i will make you know that you like no one come  i believed in your world which deep  we've a sounds of heaven on the place  it's my lord the fire is just a light  i wait or this dream loves you  baby it's a dream that love  that you're saying it much  feeling it feels  i will wait when you're too good  dream stops a reason  because we're taken apart  i've got a running on broadway  until our loves these are  you want to hurt you out today too  and they let me what a place when you wanna stop  i start freak hurting back and me and  i say it was waiting for the rest  of heaven breaking your eyes  singin' me baby  tell me  i guess if you wear    i got a party aching  it's too much we love you  see the thoughts of hearts nights in shades  i hope you couldn't run  i give a rock a chance to quite this time  should you know  i listen so see in hell but we should be wavin' it  eof | [total number of words 254]  close in the night  i left my heart without a role  some time come and he's  on the morning  i cant up the confusion  you've got to knock lisa tenderly just like a disguise  we're trying to my eyes where you've wanna talk  but you're the way you tear  push me on and shake me tomorrow  don't give it all back and take me alone  i'm hard to feel me  she'll be knowing you the best face  that you rang the shade  you can tide you free desire  and i was your baby  i could care you mess a physical feeling  i'm at that    to make you tell me when i'm falling like a woman  with a future say to me  hard to feel  yeah should you want it  tell me me tell me why  say no matter how to me  this is this my way  if i loved nobody  if who instinctively a walk back to make it  i never got  knowing the first all sweet keeps me back  to take you can't learn  to resist your and waste our sweet lord my life  repeat to hide  hoping from the nights in two surprise  the rain gods in mine  take a as prayed to me  is now it's cryin' as i reach for me to the skies  eof | [total number of words 123]  close those pillars down  we can make it started honey  you don't wanna be hearing me  like a natural thing that i do it's  maybe the way i'd be my life  i feel the rainbow  just my gaze at your eyes  side of that won't do you  no matter  i know seem to bailamos and they saw the floor  and try to make it hurts  you know you are you  it is you a dream in you  mess me me make me make you  like deep wherever i feel  when i changes beside you  do you want so your best rose    that i am  eof |

|  |  |  |
| --- | --- | --- |
| Song: honesty by billy joel, Starting word: “if” | | |
| Generated lyrics #1 | Generated lyrics #2 | Generated lyrics #3 |
| [total number of words 63]  if we relive show why and i ask  what all in love by  and i see you what you said  but i just wanna be there  you refuse for him me  always wanna another  the son of this within no    i had your laser weak  but facing a surprise of a new life  eof | [total number of words 119]  if i had my doubt you can't know  but when we you say  i couldn't think you're in love  and i love you with you  and i see us think everything will take to love  never  believe in one  and you touch so that you say  i will see my name keeps me with you  and since we're bringing    believe you  oh baby  i don't wanna do  girl i'd love you  baby  at mine  you take the one more life to come or forever  i knew yeah it's too much style your own lies  back on and love    eof | [total number of words 451]  if you want it for your best at me  even well i could told that much worse or sets it free  we think you're a sabotage  say i could knowing the music and i make us to help  i'd you do the love you know  blamin' it again  and me to me anybody be  and she's always  no one thing that we had  that all my friends  in his dreams eyes  at all music without you  reality  and if i wanna feel you  let me take it  and you can me want to each romance  i could buy  i wanna you    man i lost the clouds in the impression  love like you do  it'll soon it making  i don't refuse to go but your way  and if i'm alone  who was waiting bad  had  welcome for me me  than harmonise out  he belongs to  for the boy maybe i'm for you  i never stop  let's be your attitude  now i've tried known it together  i can you know a fool  and what can say we ain't made  another star of nothing does that it feels like the reason so much you will help  i'm alive at the night  on my baby  you wanna make all right  with me  you want to make me one more baby  i wanna be show you  my own where  so the sun is suck  closer to my heart  i'm no who loved like to show you  like to send our sign you're like it's impatiently  and that long as i say you see  and i walked out of time in summer  eyes to the street  is a box of nights of us  nothing belong to me  why you wanna be with you  and it's no way  today  what can it clean up your charms  come up now and all that you make you know how you feel for me  and this is the future is it  sailing to stay  take it  what that you are the sun  in my eyes without you  no let you hear you were shot    no one sign in time wishes clear  singing him for us  with the light strange the wind  you can't it just good on  so say as it just even relax  'cause at your at me i said  eof |

|  |  |  |
| --- | --- | --- |
| Song: lovefool by cardigans, Starting word: “dear” | | |
| Generated lyrics #1 | Generated lyrics #2 | Generated lyrics #3 |
| [total number of words 77]  dear norma gone  so i wanna stop  we know what is  listen earl come on come back like you know  we're sorry i can do  you know my eyes is a dream that  you couldn't all the way i want down  to step prove to make you tease it  just believe it's  i wanna stop and all around  but the words  'cause they're in here  eof | [total number of words 231]  dear honey i know to you  feel i do  that makes it forever  and so let the 5 walk on my own    all the warm lights on fire today  it's my body on the voice  is a way everything are your sweet heart  your latest side of the reality  honey i'm singing my own back  uh no memories come to smile  it's not too  you belong to you too  i have you  i wanna get home again  and i wanna be through  so is for you at me    hmmm why  the room you're hoping tell me starlight so you  let me hear out  no matter what makes me lose this girl  i know the way i'm warm again  you know this is your love  i feel where i sit  everybody called no shake  mess it on and ahh  born it both nothing to a na    a genie i needed from love  relax i know the matter  you want me dust the best of our lives  i don't gonna i do it  how to make those life without you  i'm tearing me let the road to can do you better  and i believe in this life  eof | [total number of words 282]  dear you gave you like you want me feel  too damn  do you need to hear this love  and i can't be there  maybe you're still  sometimes you go  just will make me it know  impossible going wrong i can't come  and i want  be left on all right  let's threw over and holiday and flow  and he's always hardly forget me  'cause that long makes every sound of us  every way of me  in his shirt  i take my dream is not comin'  gonna be your way  this is mine  love is here i'm left at the one  no one ain't what's wind for me  oh look around me  you make me love me feel like to make me up i'll hold to help  how don't it  turn me all through the wall  i don't wanna send you wanna try  betta  te quiero dich attitude  oh people come on me baby  star of beautiful are so right  because the thoughts of that way it makes the life i am  no no no  i gotta get away  but you love you breathe for us  and stay on the dawn  i believe  to tell it  leave out of that we can fly  let the highway were pacific and it seems about we know  what we planned  band of chance  feeling like you they're in the the night you can make the life dancing  oh no more time i really wanna one  eof |

|  |  |  |
| --- | --- | --- |
| Song: barbie girl by aqua, Starting word: “hiya” | | |
| Generated lyrics #1 | Generated lyrics #2 | Generated lyrics #3 |
| hiya if you're all in stranger  you gonna get me and love  you want you so much there i say  why don't do stop  say ya  oh who wouldn't do it  to make this late  if you move always tight  so remember  love start what to move away  feel where you play  sometimes how i feel it  if you want to me around  oh boy man this my best people  tearing it is it  baby don't you do you know  baby i command i wanna  let the one woman be mine  in the pretty longer fair up late  as you wave it and  tell me some more beautiful  has got a tonsils and door  i just fall at the sound  and hide  when tell as i hesitate  just the way you gave it to  my love face into me  wherever  even you'll go out of life and get hungry  breaking me go too inside  on turning me  if they disguising your lover  then honey i'm in a loss of school  on the bloom that you hear me  i can't believe at you  for me i come back  when you should run down like no no no no brace  feel like you're easy  you've got you run  if you want to feel it on the door  it's too easy to change a time  got you knowing me  or make i know confused  a long as ever far away  i knew your direction yeah and i shouldn't been just  as i couldn't believe the run  i'm apart and i'm not with the breeze  his frame on again  dee da make me come let me rock me  do you feel like those people  got me else enough tonight  eof | [total number of words 94]  hiya we are pickin'  the style of always this night  you want me move it  please feel i promised as i have not to let you crazy  i love me  and you want  don't care i'm dub of you  sadness fills my life  all i can't grow  to make you like he want to me  what you are  i turn out by my baby  you must be right but sometimes you stop and you  even wherever you feel for me  eof | [total number of words 221]  hiya all what you want is somehow you ain't want to be clear  you knowing me knowing in the love  line my mind is it  my hurts baby you're a dub  and filled without me  oh living my mind again  i feel this shame  you feel  well we were what our way  still never just to know    i think the smile  you know that i can go  and goes on the place that i will  say knowing you layin' up through the night  then wasn't when i love  i'm not fun  and i answer my own  the time is long as you could call  i'm looking to the dollar light falling down  i'll make them good enough  i can make one free home  tears is through the night  if you had no rest have not i gotta keep it  and left  at the bass inside the eyes  the pack of the livez lift you so door  you make me feel your love you  when i'm your beautiful life  to me rock you with ya  quit on his spirit alive  with a place on there everybody's never fly for you  eof |

|  |  |  |
| --- | --- | --- |
| Song: all the small things by blink 182, Starting word: “all” | | |
| Generated lyrics #1 | Generated lyrics #2 | Generated lyrics #3 |
| [total number of words 385]  all where we go  zipped away in rain  waterloo couldn't know you know i've got nobody  i just can't make our sleep till myself beg i remember ronronron  i they're give and servant  you are me free stranger  show me is rough  nothing play you  i know what you want to love me  you think i found you why  you see that makes you for dreamin'  you can't make you think it more  feels alright baby  only my prerogative  i feel kinds it's over enough  tell me what is it  one thing makes it right  don't help me renee  you're turning my own with beat  i wanna make mine  i care this thing we're sorry  then you gotta think when you facing  true can be alone  dont call me  'cause when you make me they  i love you so long  games i'm mine babe    you needed me i'm leaving  and when now we would make your brain  and hide all alone  and the neon lights lost his lookin'  my left of weak clown with you  eyes in my arms  can't make anything right down  make you do this baby  i think you gave me  like the way you're deceiving me  you're there right you wasn't  what makes you disguise  'cause you desire  just remember i said you hurt  i'm all that much  in the scene  the night makes you softly and walks  to get alone  sing you up the sky  you know i beg  bailamos  they come on all over the night  come dance when you wait  come till come through    let them it just want it  i know you know  baby i've found all i ever light  i won't beginning to believe  i have you to wait nobody won't know  but it's both words as i want nobody  one i comes is sad  'cause i left my bag  with tricks  and i seen no makes me i would step sight    eof | [total number of words 334]  all you won't tell me  'cause what your women i love you  can't tell you in the one all my own  over the meaning of time  oh my and i just wanna know  hold me  tell me tell me what he really would do  i'm there gets you  and i  can't i believe that  you're its around all over  but i'm reaching like those deep  through  in the rhythm of the charms  we must let fly knowing a magic    'cause i had a being alright  and the day is on a man  your freak filled in central  we make give me now by myself in her night  and now it is you  i am i believed  when time sometimes i cry  you're been so much afraid about me  but the sun are rush  are in the sea i cherish you for the way  because it is it  i love you at your own  looking inside  waterloo love wait for sorrow  i'll let your love you  i feel alone  turn yourself for me  everything can see  it's me that you don't like up  impossible as that special  i know i stop us  he's a reason on a new york minute  knowing the dark woman you  but now i knowing you  till how much it like they say  i planned high recognize that  loves me to the place  say crying through the door  but if am i met you  how you think i know  i think that you will know  how could you do it  this thing of you  how don't want me i left  daisy you  but you can tell what i could search  at a life with blank my dolce gone  eof | [total number of words 114]  all i know you're between me  oh i need you  my baby come on  bailamos  te quiero amor mã­o whoa  let me love again  i've got no beautiful girl  you loved to get me  let me leave your body on the arms  when i expect give a dream when we're planned  trying ya  do it is gonna be far away  it's just some chance  how couldn't follow the storm anymore  he is nice too long to anybody  'cause there's one thing guy before  heaven don't can tell me  ooh ooh  te quiero mi te quiero  eof |