LYRICS GENERATOR USING RNN

Research Proposal · May 2023

DOI: 10.13140/RG.2.2.35382.27204

4 authors, including:

Adarsh Chauhan
Maharana Institute of Professional Studies
4 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Bhavishya Arya
Maharana Pratap Engineering College
2 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Ayush Sharma
Chandigarh University
29 PUBLICATIONS **84** CITATIONS
SEE PROFILE

LYRICS GENERATOR USING RNN

Adarsh Singh

Department of Computer Science And Engineering

Maharana Pratap Engineering College, Kanpur, India

Mail:-

adarshshingchauhan.cs19mips@mpgi.edu.in

Abhishek Kashyap

Department of Computer Science And Engineering

Maharana Pratap Engineering College, Kanpur, India

Mail:-

abhishekkashyap.cs19@mpgi.edu.in

Bhavishya Arya

Department of Computer Science And Engineering

Maharana Pratap Engineering College, Kanpur, India

Mail:-

bhavishyaarya.cs19@mpgi.edu.in.

Ayush Sharma

Department of Computer Science And Engineering

Maharana Pratap Engineering College, Kanpur, India

Mail:-

ayushsharma.cs19.mpgi.edu.in

Anand Dwivedi

Professor

Department of Computer Science And Engineering

Maharana Pratap Engineering College, Kanpur,,India

Mail:- anand.dwivedi@mpgi.edu.in

Abstract-

Songs are melodic manifestations that are performed by individuals. These tunes are made altogether by both a lyricist who composes the verses and the artist who sings it. Verse writing in itself is an exceptionally selective and characterized issue. With the ever-expanding utilization of innovation and the way that they are effectively accessible

to us, makes human lives comfortable. A lyricist can often have a mind block while considering verses or may even find it difficult to get an idea. The principal reason for this exploration is to enable the lyricist to get a motivation that can assist him in making better verses. To accomplish this, a profound learning method is utilized alongside the idea of natural language processing. Specifically, bidirectional LSTM (Long short term memory) networks are used for lyric generation. The proposed framework can exceptionally create versus relying upon the information seed and the scope of words.

1. Introduction-

The verses that are composed by a musician assume a significant function in the arrangement and formation of the song. It gives life to the music and makes the song whole and complete. But the work behind forming the lyrics is often

tiresome and time-consuming. Additionally, if there is a writer's block or no inspiration, then the lyric writing process is delayed further. Machine learning has already done wonders in backing up humans by automating tasks completely or partially. In this case, also machine learning can be used as a tool to help the generation of lyrics.

The proposed system aims at generating lyrics according to the seed input which is given by the user. In this way, the lyrics will be generated following the idea that the writer wants. Backed by deep learning and Natural Language Processing (NLP) concepts, this system is trained on a unique dataset of 'romantic' English songs. Long Short Term Memory networks (LSTM) [1], an uncommon sort of recurrent neural network (RNN) are utilized for the learning process. These LSTM networks are equipped for recalling data for a more drawn out period because of the presence of cell states.The intuition is that the generation of music relies upon the arrangement of input words we feed and subsequently the next word generated must relate to it,

therefore the utilization of the LSTM model is favored over the ordinary RNN model. Before the data can be given to the LSTM networks for learning, it is pre-processed by using NLP techniques. The insight is that since the lyrics consist of the words and strings, we may need to preprocess them using NLP to increase the performance. Specifically the bidirectional LSTM networks are used since the efficiency is better than the other models it's compared with, namely LSTM and gated recurrent unit (GRU) networks.

2. Implementation

The implementation is demonstrated in steps and shown in Fig.2:

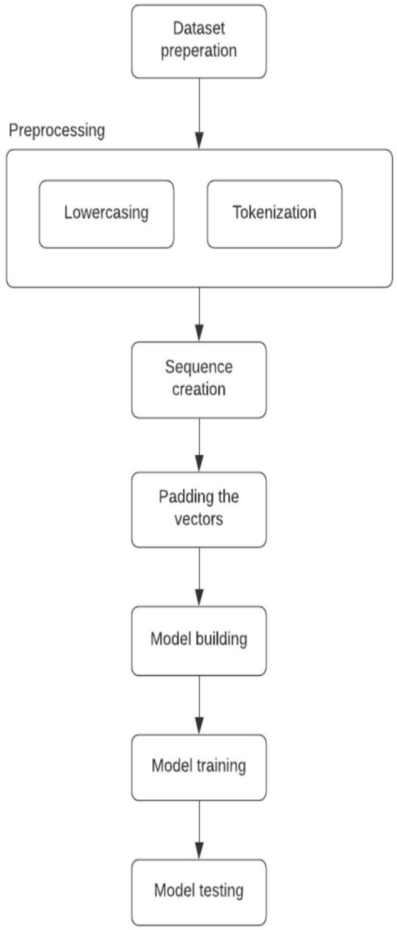


Fig.2. Flowchart of the implementation

Dataset preparation

The dataset is unique and customized according to the use case. It consists of lyrics of English songs scraped from the web. The total number of words in the dataset is approximately 13k.

3. Preprocessing-

The dataset is preprocessed to maintain consistency in data which will help further during the training of the model. The pre-processing techniques used are as follows:

Lowercasing The dataset contains both capital and small letters. Since the model will take the same word with a difference in capitalization as distinctive, it will create confusion. For instance, 'The' and 'the' is unique for the model to manage and may not anticipate the similarity on encountering it. Hence, making all the letters lowercase would wipe out this problem and improve the results.

Tokenization The data to be used in the model should be of numeric structure for further analysis and classification. This is accomplished by the process of 'tokenization'. This helps by making the numeric tokens of each word present in the corpus.

Sequence creation

The next step is to turn the sentences into lists of values based on the tokens generated by the tokenizer. This is significant as the model will anticipate the next word on feeding the set of input words, we need to train it on a set of words with the next word as output to our model[6].

Padding the vectors

In padding, the list of sentences[7] has been padded out into a matrix of the same length. This is achieved by putting the appropriate number of zeros before the sentence list. The matrix width is kept the same as the width of the longest sentence to make each vector of the same length.

Model building

Now, since all the required preprocessing is done, the next step is to build the model for training. The layers that are used for model building are as follows:[9]

Embedding layer It is the first hidden layer of the neural network. The intuition is that words and related words are bunched as vectors in a multidimensional space. The input

and output dimension along with the length is specified. For this layer, the yield is a 2D vector with one embedding for each word in the input sequence of words.

LSTM layer LSTM layer and bidirectional LSTM layer is used. The bidirectional layer helps in providing additional context to the network and results in a faster learning process. The bidirectional LSTM trains two LSTMs on the input sequence. The first one is the input sequence the way it is and the second on a switched duplicate of the input sequence.

Dropout layer This layer is introduced to reduce overfitting. At each update of the training phase, it works by arbitrarily setting the outgoing edges of hidden units to 0.

Dense layer This layer is fully connected. All the neurons in a layer are associated with those in the following layer. It is followed by the activation function which in this case is "softmax".

4. Model training-

The model is compiled using the loss function "categorical cross-entropy" and optimized using "adam's" optimizer [4]. Then it is trained on 100 epochs by fitting the parameters.Another research named Tra-la-Lyrics 2.0 was done by the author Gon¸calo Oliveira, H. The exploration referenced a framework for the automatic generation of song lyrics on a semantic area. It is the improvement of the original Tra-la-Lyrics wherein the content is created which is characterized by at least one seed word. To gauge the advancement, the mood, the rhymes, and the semantic intelligence in verses delivered by the first Tra-la-Lyrics were investigated and contrasted with the verses created by the new launch of this framework, named Tra-la-Lyrics 2.0. The assessment demonstrated that, in the sections by the new system, words have higher semantic relationship among them and with the given seeds, [11] while the rhythm is so far composed and rhymes are accessible. A study was led to affirm the aftereffects of the improvement of the verses by Tra-la-Lyrics 2.0 [2]. Furthermore, the author's Eric Malmi, et al., proposed DopeLearning. It explains the rap lyrics generation using a computational approach. They build up an expectation model to recognize the following line of the current verses from a lot of applicant next lines. The model is based on two artificial intelligence procedures. They are the RankSVM algorithm and a deep neural network [8] model with a novel structure. Results show that the forecast model can recognize the valid next line among 299 erratically picked lines with a precision of 17%, i.e., more than 50 times more likely than by discretionary. They additionally utilize the forecast model to join lines from existing tunes, creating verses with rhyme and significance. An assessment of the delivered verses shows that as far as quantitative rhyme thickness, the technique beats the best human rappers by 21% [3].

5. Model testing

The model is tested with a set of input words to begin the lyrics with and word limit. The testing is to be carried out after preprocessing of the input and then parsing it through a function that predicts the next possible word for that phrase. This prediction runs in a loop to give out the chain of words to form lyrics.

The model used has 25 input variables and it is formed by firstly 'Embedding layers' with 160 neurons, the second layer is 'Bidirectional LSTM' with 400 neurons followed by a dropout layer of 0.2. Third is a LSTM layer with 100 neurons, followed by two dense layers with activation function 'relu' and 'softmax' respectively. The final dense layer of 23000 class is then followed by a compilation layer. A total set of trainable parameters (1,298) was compiled using 'adam' optimizer and 'categorical crossentropy' as the loss function. The layers of the CNN architecture can be seen in Fig.3.

6. Results-

The dataset is trained on three models. These models are LSTM network, bidirectional network and GRU network model. The best results were obtained from the bidirectional LSTM model. This model is trained on approximately 18 lakh parameters for 100 epochs. The accuracy achieved is 82.6% and the loss is 0.40. The accuracy and the loss results for the models are shown in Fig.4. Since the

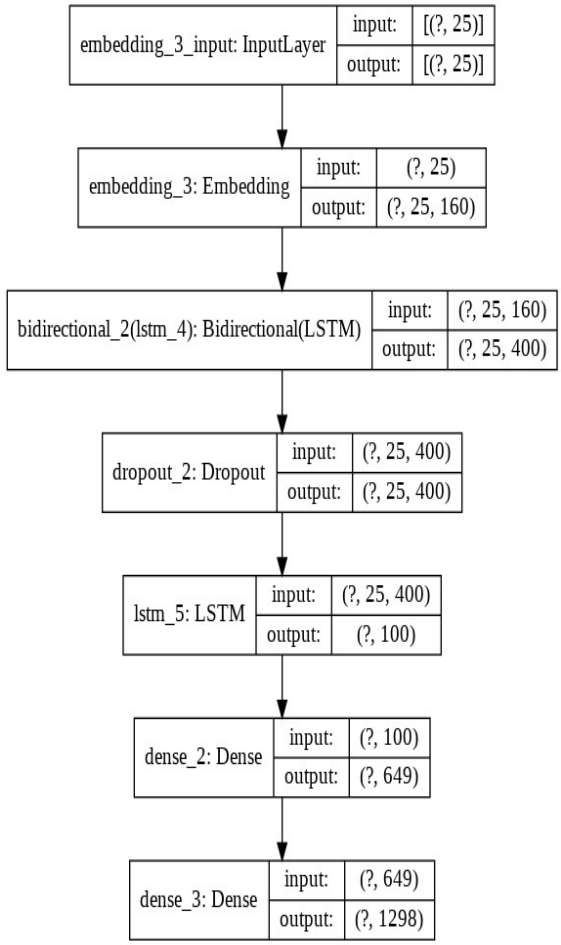


Fig.3. Layers of the model

accuracy of all the models have very less difference, for more validity the output for different models are shown in Fig.5 for a sample input. The predicted lyrics for bidirectional LSTM model made more sense and had better sentence formation.

7. Conclusion-

The lyrics inducer model was successfully built using natural language processing and deep learning. Bidirectional LSTM network has proved to efficiently remember the connection between the words which helps in predicting better song lyrics. The Training accuracy of 82% achieved can produce decently meaningful results according to the given seed input.

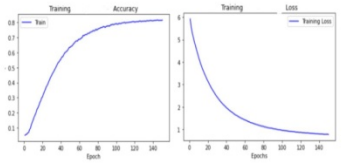
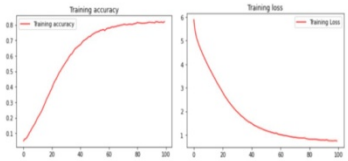
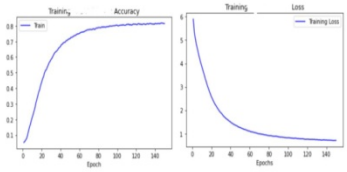
Sr. No	Algorithm	Training accuracy	Training loss	Plots
1	LSTM	81.50%	0.76	
2	Bidirectional LSTM	82.60%	0.43	
3	GRU	81.14%	0.71	

Fig.4. The training accuracy and loss of LSTM,Bidirectional LSTM, and GRU

Future scope of the model could be to add more data to the corpus and train the model on more epochs. Furthermore, models can be built for separate genres for more genre-specific lyrics.

Input seed	Word limit	Prediction		
		LSTM	Bidirectional LSTM	GRU
Singing alone	50	Singing alone we make of what we met you fall my future in your temper is a promise me and now a bore is already gone gone gone gone gone gone already gone this gone love gone already indeed now already gone already gone already gone how to gone already gone already	Singing alone just about to get you crazy never give you everything so heavy come on the touch floor right there off doheny yeah every mood need you brr brr brr mind you need that lead eyes maybes is key for seventy minutes came like yeah every sin alone hey hey	Singing alone in my sleepless solitude tonight me laugh like i hates but i see my mind up yet we'll mend your heart and so big ground friend out missing wild diamond free free just water baby still it's no to i'm to give you after all these years no i'm not
Her hair	50	Her hair her hair falls perfectly without her trying to believe that now you let me go never let me go never let me gone beneath gone this high when you're already me gone love gone gone gone gone gone gone gone gone gone gone gone gone gone you already gone	Her hair curves i'm ghost just go unnecessary came south in ink and bleach like the '80's came shoot note we been trappin' like the '80's came shoot forever like me like through breaking but fm superhuman at the lot on the sea big jo ek jamati pourin' don't give you all	Her hair her hair falls perfectly without her trying at us to do us like i kissed you home me yet it's gone the pain truth just a little so used your eyes crazy crazy for you baby if you see me how to bend without the world caving in your faith
Toosie Slide	40	Toosie slide baby you got me like ah woo ah ow out to feeling what i'm feeling time with you baby i'm in love with you already gone this gone love to let it gone gone gone what you want to gone	Toosie slide you talking 'bout me i don't see a shade mm like to get me up in my bag to let me up and throw a tantrum all you just like to let me take you down i'm dwayne carter no	Toosie slide the whole world stops and stares for a while thing i can escape the stars they see the words of me i'll be your voice put a lips just kids girl lot just kids bit so dreaming wild love beautiful
The box	30	The box world is bright all right here all along in love babe again and again and again and again and again and again gone gone gone what gone this full gone	The box that tight they comfort in a man and bleach like monopoly baby say style I can carry dont do it is shine for me i don't know you care bout	The box whole world stops and stares for a while thing i can escape the stars they see the words of me i'll be your voice put a lips just kids girl

Fig.5. Input and output results of LSTM,Bidirectional LSTM, and GRU

8. References-

1) (n.d.). Institute of Bioinformaticspage. https://www.bioinf.jku.at/publications/older/2604.pdf

2) Gon¸calo Oliveira, H. (2015). Tra-la-Lyrics 2.0: Automatic Generation of Song Lyricson a Semantic Domain, Journal of Artificial General Intelligence, 6(1), 87-110. doi:

3) https://doi.org/10.1515/jagi-2015-0005

4) Eric Malmi, Pyry Takala, Hannu Toivonen, Tapani Raiko, and Aristides Gionis.2016. DopeLearning: A Computational Approach to Rap Lyrics Generation. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 195–204. DOI:https://doi.org/10.1145/2939672.2939679

5) Diederik P. Kingma and Jimmy Lei Ba. Adam : A method for stochastic optimization. 2014. arXiv:1412.6980v9

6) "Generating Sequences with Recurrent Neural Networks" (Alex Graves) - This paper provides a detailed explanation of how to use RNNs, specifically the Long Short-Term Memory (LSTM) variant, for sequence generation tasks. It covers the basics of RNNs and provides insights into training and sampling from the model.

7) "The Unreasonable Effectiveness of Recurrent Neural Networks" (Andrej Karpathy) - This blog post by Andrej Karpathy demonstrates the power of RNNs for generating text, including song lyrics. It provides a step-by-step guide on training an RNN-based language model using character-level prediction and showcases interesting results.

8) "Generating Text with Recurrent Neural Networks" (Ilya Sutskever, et al.) - In this research paper, the authors present a technique for training RNNs to generate text. They discuss the challenges of long-range dependencies in language modeling and propose an approach using LSTM networks. The paper includes experiments on generating Shakespearean text and demonstrates the capabilities of RNNs for creative text generation.

9) "DeepBach: A Steerable Model for Bach Chorales Generation" (Gaëtan Hadjeres, et al.) - While this paper focuses on generating classical music, it introduces a methodology that can be adapted to lyrics generation. It explores the use of deep learning techniques, including RNNs, to capture musical structure and generate coherent compositions. The concepts and techniques presented can inspire the development of a lyrics generator.

10) "Lyrics Generation with Recurrent Neural Networks in TensorFlow" (Jason Brownlee) - This tutorial provides a step-by-step walkthrough of implementing a lyrics generator using RNNs with TensorFlow. It covers data preparation, model architecture, training, and generating lyrics. The tutorial also includes code examples to help you get started quickly.

11) "Neural Text Generation: A Practical Guide" (Alec Radford) - In this blog post, Alec Radford, one of the authors of the GPT model, provides insights into training text generation models, including lyrics generation. The post covers techniques such as teacher forcing, beam search, and temperature sampling, which are crucial for improving the quality and diversity of generated lyrics.

12) "Generating Text via Adversarial Training" (Zhang, et al.) - This research paper introduces a framework for text generation using generative adversarial networks (GANs). It explores the idea of using RNNs as the generator component of the GAN and discusses training strategies and evaluation methods. While the paper focuses on general text generation, the techniques and concepts can be applied to lyrics generation as well.

13) "Neural Network Methods in Natural Language Processing" (Yoav Goldberg) - This book covers various neural network techniques for natural language processing tasks, including text generation. It provides a comprehensive introduction to RNNs, LSTMs, and other related architectures. Chapter 9 specifically focuses on language generation models, which can be adapted for lyrics generation.

14) "Magenta" by Google - Magenta is an open-source project by Google that focuses on the intersection of machine learning and creativity, including music and lyrics generation. The project provides various tools, models, and datasets for generating music and lyrics using RNNs and other deep learning techniques. Exploring the Magenta project can give you valuable insights and resources for your lyrics generation project.