**Song & Lyrics Generator**

(Deep Learning and NLP Libraries)

### 

### 

J Component Project Review 3 for the course-

CSE4022 Natural Language Processing

Submitted to-

### Prof. Ilakiyaselvan N

By-

Bhavesh Kumar - 19BCE1313

Kshitij Upadhyay - 19BCE1305



**Abstract**

Natural language processing is one of the most attractive and difficult streams or areas in machine learning or artificial intelligence. It is different from computer vision and other machine learning tasks as NLP doesn’t convey meaning through any tangible manifestation. By the virtue of deep learning, NLP has achieved tremendous progress in keyword search, machine translation, semantic analysis, etc.

Neural network-based language models have shown promise in producing unique lengthy-shape prose content from the minimal initial textual content. This work has been further extended to text generation.

Text generation is one of the most common examples of applied Machine Learning (ML). The constant evolution of algorithms and the huge amount of text data available allows us to train models that try to capture context or imitate previous work from others. In this way, ML has been used to do everything from writing journalistic articles to short stories to entire novels. Unfortunately, many of them are obviously poor examples of writing mainly because the ML routines often generate nonsensical sentences and generate long-form documents with poor structural coherence.

Interestingly, these shortcomings work in ML’s favor when it comes to writing poetry, and by corollary, lyrics since they’re more focused on generating an emotional response rather than an intellectual one in the reader/listener. Given a set of input keywords, and ML-driven lyric generator will produce a poem that won’t win any awards, but will likely evoke a response from the reader, even if it’s just to laugh at its absurdity.

In this project, I am trying to make a model which will generate lyrics based on the seed text it receives. I will be making a model for world-level prediction, by pre-processing and embedding the text before passing it into the model before both trainings the model and predicting on test data.

Introduction

We will be trying to make a word prediction type of system as it has many applications like dealing with queries as a chatbot, writing emails or other documents, or texting and chatting.

Word Prediction has many use cases, and its use are vast. Moreover, Lyrics/ Poetry generation can also be helpful for various kinds of assignments and professionally.

We will try to use deep learning and to make LSTM models for the prediction of words. We will also be using Python and Jupyter, and Keras Library (for implementing neural networks). We will also be trying to use TensorFlow for training the model and predicting.

First, we will train the model on 200/250/300 epochs. It will take some time, and then we will use that and save that model by using that we can predict the lines.

We are using the iconic William Shakespeare poetry to begin within our model, but after completion of the model, we can also use other text like perhaps dialogues of evergreen Friends(a popular TV show)

**Literature Survey**

Lyrics Generation using LSTM has been a viable option in the past. [1] attempted to generate lyrics as a creator of lyrics for a specific artist. However, their model was limited in generating lyrics for a genre, as it was trained on a specific artist. Additionally, [1] implemented rhyming by training their model with sets of lyrics in which corresponding rhyming words were noted.

[2] created a model that produces entire lyrics for a given input melody. Pairing lyrics and the Deep Learning in Musical Lyric Generation: An LSTM-Based Approach. This paper explores the capability of deep learning to generate lyrics for a designated musical genre. They employed the LSTM network to produce lyrics for a specific genre given an input sample lyric. The problem they model suffered as they were only able to generate lyrics in 1 genre alone hence no diversity.

[3] manually analyzed correlations among melodies, beats, and syllables using 42 Portuguese songs and propose a set of heuristic rules for lyrics generation. However, it captures only phonological aspects of melody-lyrics correlations and can generate a small fragment of lyrics (not an entire lyrics) for a given piece of melody.

[4] attempt to induce a statistical model for generating melodic Tamil lyrics from melody-lyrics parallel data using only ten songs. It suffered from a severe shortage of data and fails to conduct empirical experiments.

[5] The program consisted of a database of about 40,000 existing rap lyrics. A new lyric is then generated from the words and verses found in the existing lyrics. A linear-interpolated trigram model approach was used to make the lyrics. However, the result was rated as lacking flow. Therefore, they shift to a quad ram model. They also implemented a database containing rhyming words from two different sentences. In this way, they generated sentences that rhyme with each other. Lastly, all the sentences were pieced together according to the song structure and layout. The generator worked fine but the content of the lyrics was not meaningful and did not relate to a specific theme.

Wishful Automatic Spanish Poet was the first poem generating program which combined natural language generation techniques with artificial intelligence [6]. It is a system that obtains inputs from users, and these were used as seeds. The system is based on a forward reasoning ruled-based system. The results obtained were evaluated as being poor and not very efficient.

[7] experimented on rhyme and style features to classify and process lyrics. They used a group of words together with part-of-speech tagging and other statistical features for processing lyrics. A rhyme is actually two words that when spelled sound similar. This feature is generally used for words at the end verses. A proper evaluation of the proposed method was not done.

[8] uses CFGs to generate lyrics. This system does not make use of a predefined template of words that just adds the user inputs into pre-defined places. Instead, sentences in the lyrics are constructed based on n-sentence grammar. A random combination of word bi-grams and tri-grams is used to form a verse of a certain length. The problem faced is that the lyrics generated are not many logical dues to various ambiguities in NLP and hence don’t always produce good results.

[9] experimented on the use of lyrics to automatically identify and classify music and to determine artist similarity. Song lyrics were collected from various sources on the web. Various techniques such as the PLSA (Probabilistic Latent Semantic Analysis) and the k-means clustering methods were used to analyze the content and semantics of the lyrics. The evaluation was carried out by matching the system together with another audio system to check their similarity. Both techniques used had their pros and cons. Hence, a combination of the two techniques could prove to be much better but this was left as future works.

An encoder-decoder-based RNN sequential model for lyrics-conditional melody generation for Chinese pop songs is presented in [10]. The sequential model called Songwriter consists of two encoders and one hierarchical decoder. The encoders are designed to encode the lyric syllables and context melody of the prior generated melody. The hierarchical decoder is designed to decode the note attributes such as pitch, duration, and syllable note alignment labels since most of the syllables in Chinese songs had more than one note.

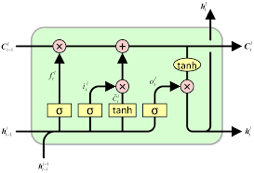
**Problem Statement**

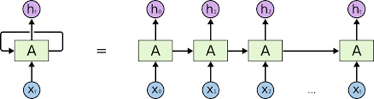
From the papers above, we can observe that logical lyrics generation and good accuracy have been a problem. We will be using a logical and diverse dataset and will make an optimal model in an effort to counter the above problem and provide satisfactory results.

**Architecture**

From previous research articles, various articles available on various websites ex- medium, it is evident that the LSTM layer works well on text data primarily because LSTM work very well on sequential data. It not only passes its output to the neurons in the next layers but also passes them to adjacent neurons in the same layer. They have internal mechanisms called **gates** that can regulate the flow of information. These gates can learn which data in a sequence is important to keep or throw away. By doing that, it can pass relevant information down the long chain of sequences to make predictions.

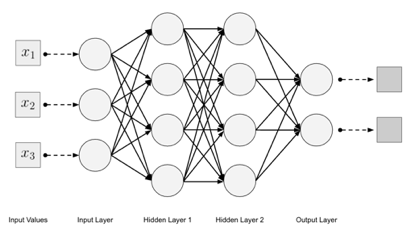
**LSTM Cell architecture**

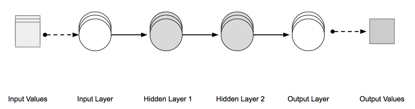




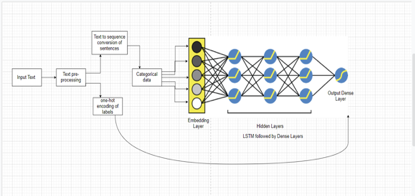
**Project Architecture –**

1. **General DL Model-**

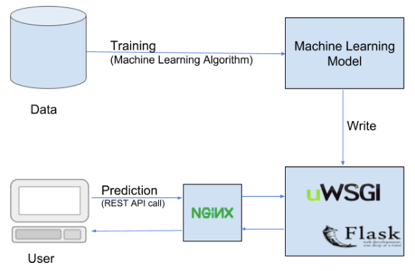




**Song Lyrics Generator (Project) –**

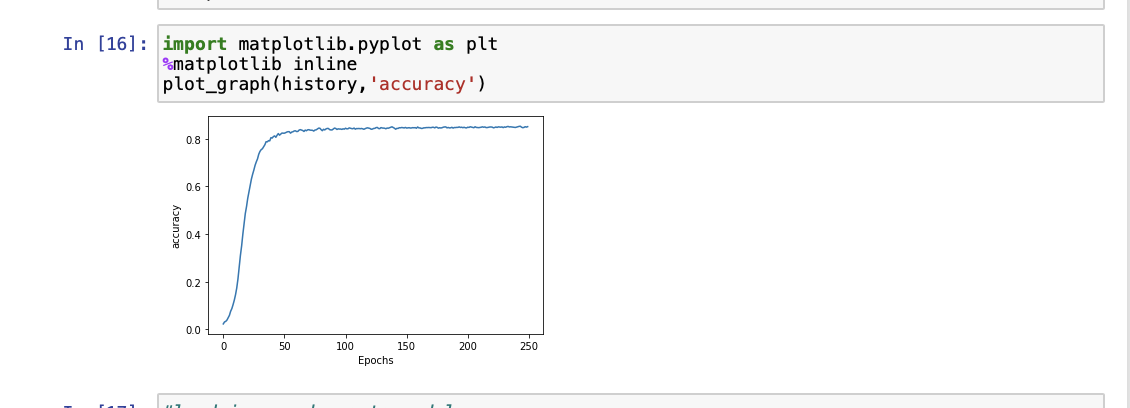


1. **Final predicted Architecture after integrating UI through Flask**



We did text preprocessing by tokenizing, sequencing, and labeling them and then we applied the lstm model and added some layers like dropout layer with the value 0.2 and then we trained our model for 250 epochs and got the accuracy of approx 85%.

**Plotting the accuracy:**

****

**OUTPUTS**

Following are some output examples where we have given some input words as seed text and have given the number of words we want our model to predict and then we get the output for the same.

Input words are “This is life” and we want the next 4 words to be predicted

The output which we got is “This is life on me and doth”

Input words are “The expense of spirit” and we want the next 4 words to be predicted

The output which we got is “The expense of spirit is a waste of”

Input words are “That thereby beauty’s ” and we want the next 4 words to be predicted

The output which we got is “That thereby beauty’s rose might never die”

Input words are “That thereby beauty’s ” and we want the next 6 words to be predicted

The output which we got is “That thereby beauty’s rose might never die not show”

Input words are “That thereby beauty’s ” and we want the next 2 words to be predicted

The output which we got is “That thereby beauty’s rose might”

Input words are “wrath of the life ” and we want the next 4 words to be predicted

The output which we got is “wrath of the life that thou shouldst bear”

To conclude we can see that as per the initial words or lyrics our model predicts the further words and the accuracy of our model is more than 85%.

**Literature cited**

1. Potash, P., Romanov, A., Rumshisky, A:. GhostWriter: Using an LSTM for Automatic Rap Lyric Generation. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. (2015)
2. Watanabe, K., Matsubayashi, Y., Fukayama, S., Goto, M., Inui, K., Nakano, T.: A melody-conditioned lyrics language model. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), vol. 1, pp. 163–172 (2018)
3. Hugo Gonc¸alo Oliveira. 2015. Tra-la-lyrics 2.0: Automatic generation of song lyrics on a semantic domain. Journal of Artificial General Intelligence 6(1):87–110.
4. Ananth Ramakrishnan A, Sankar Kuppan, and Sobha Lalitha Devi. 2009. Automatic generation of Tamil lyrics for melodies. In Proceedings of the Workshop on Computational Approaches to Linguistic Creativity. pages 40–46.
5. Nguyen H. and Sa B., (2009). Rap Lyrics Generator. Unpublished.
6. Manurung H., Ritchie G. and Thompson H. (2000). Towards a Computational Model of Poetry Generation. In: Proceedings of AISB Symposium on Creative and Cultural Aspects and Applications of AI and Cognitive Science, pp. 79-86, April 2000, Birmingham.
7. Mayer R., Neumayer R. and Rauber A. (2008). Rhyme and Style Features for Musical Genre Classification by Song Lyrics. In: ISMIR 2008, 9th International Conference on Music Information Retrieval, 14-18 September, 2008, Drexel University, Philadelphia, PA, USA.
8. Automated Generation of Song Lyrics using CFGs. Sameerchand Pudaruth, Sandiana Amourdon and Joey Anseline, Department of Computer Science and Engineering, University of Mauritius, Réduit, Mauritius.
9. Logan B., Kositsky A. and Moreno P. (2004). Semantic Analysis of Song Lyrics. In: Proceedings of the IEEE International Conference on Multimedia and Expo, pp. 817-830, 30 June 2004, Taipei City, China.
10. 10. H. Bao, S. Huang, F. Wei, L. Cui, Y. Wu, C. Tan, S. Piao, and M. Zhou, “Neural melody composition from lyrics,” arXiv preprint arXiv:1809.04318, 2018.

Appendix :

This is the dataset that we used

