International Journal of Engineering & Technology, 7 (2.8) (2018) 677-679

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper



A survey on automatic music generation

Shubham Jain*[1], A. Pandian[2]

[1]Student, [2]Associate Professor,
[1,2]Dept. of Computer Science & Technology, SRM Institute of Science & Technology, Chennai, India
**Corresponding Author Email: shubhjain.helloworld@gmail.com*

Abstract

Just like you should not watch a foreign language movie without its subtitles, identically you should not listen to music without its lyrics Music lyrics are words that combine to produce a song in harmony. Usually the music lyrics that we normally listen to are human written and no machine involvement is present. Writing music has never been a easy task, lot of challenges are involved in this because the music lyrics need to be meaningful and at the same time it needs to be in harmony and sync with the music being played over it . They are written by great artist who have been writing music lyrics for years. This project tries to automate the process of music lyrics generation using computer program which we produce lyrics and reduce the burden on human skills and can generate new music lyrics and a very faster rate than humans ever can. This project also aims toward the merge of human and artificial intelligence.

Keywords: *Music Lyrics; Artificial Intelligence; Markov Model; LSTM.*

1. Introduction

Writing music lyrics is an art and requires tremendous efforts to sync each word to another and produce a meaning out of it. Music lyrics are what defines the tone of the musical tunes, so music lyrics is a platform for the creative world to showcase their talent to the world. The task that is very difficult for humans to perform might look even more difficult for a machine to do, well this project is all about the machine generated music lyrics. My focus is introduce a digital machine to generate new music lyrics, what if we can make the machine to understand the lyrics and then even write new music lyrics for us. Several scholars have been already trying and understanding the music lyrics auto generation by machine learning models, but very few have successfully been able to achieve the target of generating new music lyrics that have a good sync within the words itself and have a semantic structure out of the generated music lyrics.

My project on this topic aims to generate a meaningful music lyrics given a word description about the song to be composed. According to today's situation lyricists are given a mood or the going scene from the director and the people composing music lyrics write the entire song lyrics on their own out of their creative idea and emotions. They even try to alter the words from old poems and song lyrics by adding few new words to it and changing the overall emotion of the song. The model created in my project generate new music lyrics by using the old music lyrics as the input for the programs, the more the input data better the quality of the music lyrics. The model tries to match and combine words of one music song lyrics with the lyrics of other songs and a new music lyrics are generated in this process.

The biggest challenge faced in doing this project is the semantic link between the words in the music lyrics at times when the new music lyrics are generated it may or may not have a meaning out of it, so to avoid this a huge dataset of music lyrics are necessary.

More the input data better the output will be because the model will get more words to chose from during generating of new music lyrics.

2. Survey Literature

Since there is a lot of creative potential in the automatic music lyrics generation, many researchers have been made to try and understand the auto generation of music lyrics. Using artificial intelligence into music has a rich history dating since late 1980s. Various machine learning algorithms have be consider for automatic music generation ranging from neural networks to expert systems. But since lack of data set during the early research of auto music generation always restricted the further process, but now because huge data is available with us this task has got easier than before.

Several attempts have already been made but one such attempt in particular have got some great results recently. Pair of researchers from Hong Kong University, decided to download lyrics of 52000 rap songs from the internet and then train the model that they have created to give a novel song. Their model could identify the rhyme scheme.

Their model generates the feature vectors for larger bispans, recursively from smaller bispans using both input and output language contexts simultaneously. So when two smaller bispans combine, they need to compress two feature vectors of dimension d corresponding to both the bispans, to one feature vector of dimension d. First bit of the input is set depending on whether they combine in straight or inverted order. The vectors are normalized at each stage to prevent the model from reaching a degenerate optimum of arbitrarily small vectors. For example, their program generates output that rhyme words not just with the end of the sentence, but also in the middle of challenges, which indicates that their model captures the structural associations more preciously than the phrase-based model.

Another researcher from San Jose, United States of America published a paper where he introduced ALYSIA, which is a automated music lyrics writing program. ALYSIA was created using a very famous machine learning model called Random Forest. He showed how he used ALYSIA to generate original pop style music lyrics song successfully and produced them as well.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

He discussed how he used the glory of ALYSIA for generating rhyme and pitch predictions. Without any much of the modification required and with no human involvement, ALYSIA can be used to generate complete song by just selecting the very starting suggestion for each line of music lyrics. Two models where created first for the note duration prediction and another for scale degree of the note prediction. The first model was referred as Rhyme model and the second was referred as melody model. The most crucial part of the model created is the features set that were used, because the aim was to create features that will help to understand the structure of the music lyrics and its compositions.
The input that was provided to the model was melodies for the text that was given by the users, their training sets was composed of the sentences in the verses of music and not the random melodies. 75% of the data was randomly split for the training purpose by stratified sampling and the remaining data was used for the testing purpose.

Table 1. Accuracy chart

MODEL	ACCURACY
Rhyme model	86.79%
Melody model	72.28%

Total of three music lyrics were composed using the ALYSIA, the first and the second music lyrics depended on the writer and the last one was borrowed from a 1917 music song. All the music lyrics were produced and are available on http://bit.ly/2eQHado . The creators of the model are regularly experimenting with the model's potential for writing new music lyrics in different genres.
Recently a group of researchers from Finland have published a paper titled: Dope Learning. It is a computational approach towards music lyrics generation. They created a algorithm called DeepBeat. The algorithm's aim was to predict the most relevant line for the previous lines fed to the model as the query, this can also be seen as a information retrieval model. They used a famous machine learning algorithm to create their model called as a ranked Support Vector Machine, which was partially powered by a deep neural network and they fed it with a dataset of all the songs from the top 100 english-speaking rap artists.
The program that was created helps in extracting three kinds of features:
1.Rhyme features.
2.Structural features.
3.Semantic features.
Out of all those three metrics the, semantic similarity was the one that required the use of a deep recurrent neural network and the neural net did what it does best it found vector representations of words, lines and groups of lines. Once the features were calculated they were input into the ranked SVM model. SVM model is type of broad linear classifier and the SVM eventually learnt to predict the next line once it was trained on the input features.
They evaluated the model's prediction performance by amounting how accurately the model predicts next line of the a given rap song. 82% of accuracy was achieved, which means the next line that was generated was 82% different from the random line of the input music lyrics line. They also wanted to find a way to quantify how good their algorithm was compared to human experts and so they calculated something called rhyme density, which is the average length the longest Rhyme per word using rhyme density as a metric. They found that the algorithm's generated lyrics had a 21% higher density than the most rhyme dense human artists on their lists.
In another work by a famous group of researchers at University of Massachusetts Lowell ,US they have shown the accuracy and the efficiency that is achieved by using s LSTM model for the new music lyrics generation. They prove this by comparing the LSTM model with a simple model called n-gram model. They built a model that requires no templates to generate lyrics at the same time it should also be able to produce full verses of the music song unlike the other existing system that are capable of producing only single lines of music lyrics. Their model is also capable of creating new music lyrics that matched the musical style of the artist's inputs. It was achieved by understanding vocabulary and rhythmic style of the artists. This will enable the model to create similar style music lyrics yet a new one.
The model created was fed with songs collected from 14 artists from the internet. For training purpose they used 219 verses with minimum of 175 words in each of the verse. They built the LSTM on top of Theano dependencies in Python. Both the inputs as well as the outputs were equal in vocabulary size. They trained their network using a Tesla GPU on a single workstation and in order to avoid the situation of disappearing of gradient problem during the training of the networks they set the gradient parameters in the range [-1,1]. When fed with a the inputs of words as (wk,...,w0) and to determine P(wk+1|wk,...,w0; E,Φ), where Φ stands for the parameters of the network created.
They defined the probability by:

$$P(wk+1|wk,...,w0;E,\Phi) = f(x,s) \quad (1)$$

Later they even used the generated rhyme density of 0.34, which they used to check the similarity needed to achieve the target rhyme density.

Table 2. Similarity chart

Model	Similarity
N-gram model	1.28
LSTM	0.59

By looking at this numbers they proved that LSTM model very effectively outperforms the n-gram model, when generating new music lyrics.
Another group of young researchers from Anna University, Chennai have proposed a very idea of generating music lyrics using Dijkstra's shortest path algorithm. The model developed is used to predict pattern for a new melody to generate optimal sequence of syllable, which is then fed to the lyrics generating module that come up with new meaningful phrases and verses that matches with the syllabic pattern.
The model used a corpus that consisted of 10 melodious as the training data to understand the patterns of the syllable. Tamil, their target language was generating lyrics. Tamil being a phonetic language has a one on one relationship between grapheme and phonemes. The approach which they used to generate new music lyrics for any given input melody is a two stag method. The first is to parse the input melody and output a series of syllable pattern in KNM representation scheme and the second stag is combing the syllable pattern with the words from the corpus.
For any given input, for example 'ABAB BAB ABA' , they constructed a directed graph with the list of words that matches the patterns extracted.

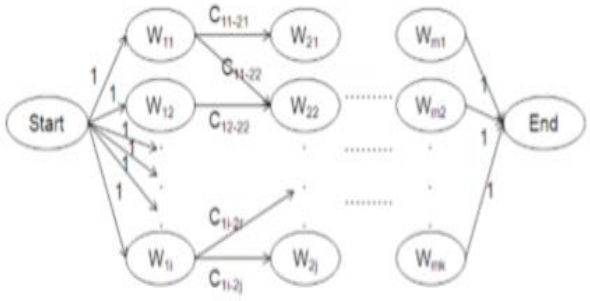


Fig. 1: Paths possible.

Since the algorithm picks the shortest path among all the paths that are available in the graph. So the path that has the least cost function value is extracted. One of the limitation of using this is that at times the algorithm will always generate the same phrase for it.

3. Challenges

The biggest challenge while doing any artificial intelligence project is that the accuracy is never 100% and therefore the outputs generated are not always true and correct, they are mere predictions that are formulated, which may or may not result into a true output. The output generated in this project may not always have a proper semantic links between the words of the newly generated music lyrics.
The second biggest challenge is of crisp data that is required to fed into the model that is created. Lot of data that is available today is raw and need a lot of preprocessing before using them. Lack of computational power also is treated as a limitation to do any kind of machine learning project.

4. Conclusion

Automatic music lyrics generation have always been an interesting topic for researchers across globe. All the proposed idea in above literature have a crucial advantages of their own as well as limitations as well. Therefore no system is perfect as no human ever is. The Markov model that is used in this project has a python implementation. To built the auto music lyrics generation , only two machine learning libraries are used, one to generate random numbers and another to deal with the text format.

5. Future Work

In future I will keep on experimenting with new datasets and newly created machine learning models that can generate new music lyrics by learning from the inputs fed into it. I will even experiment with new regional Languages of India. Focus will also be on integrating a speech synthesiser that will give the algorithm a voice.

References

[1] Eric Malmi, Pyry Takala, Hannu Toivonen, Tapani Raiko, and Aristides Gionis. DopeLearning: A Computational Approach to Rap Lyrics Generation, arXiv :1505.04771v1 [cs.LG], 18 May 2015.

[2] Peter Potash, Alexey Romanov, Anna Rumshisky. GhostWriter: Using an LSTM for Automatic Rap Lyric Generation, proceedings of conference on Empirical Methods in NLP, 2015

[3] Margareta Ackerman and David Loker. Algorithmic Song writing with ALYSIA, arXiv:1612.01058v1 [cs.AI], 4 Dec 2016.

[4] Jukka M. Toivanen and Hannu Toivonen and Alessandro Valitutti. Automatical Composition of Lyrical Songs, in The Fourth International Conference on Computational Creativity, 2013.

[5] Automatic generation of song lyrics on a semantic domain. Journal of Artificial General Intelligence, 6(1):87–110, 2015.

[6] Dekai Wu and Karteek Addanki, Learning to Rap Battle with Bilingual Recursive Neural Networks, proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

[7] Sravana Reddy and Kevin Knight. 2011. Unsupervised discovery of rhyme schemes. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2, pages 77–82.

[8] Toivanen, J. M.; Toivonen, H.; Valitutti, A.; and Gross, O. 2012. Corpus-based generation of content and form in poetry. In International Conference on Computational Creativity, 175–179

[9] Kristine Monteith, Tony Martinez, and DanVentura. Automatic Generation of Melodic Accompaniments for Lyrics. Proceedings of the Third International Conference on Computational Creativity ,2012.

[10] Diveesh Singh, Helen Jiang, MindyYang. Automatic Lyrics Transcription by Separating Vocals from Background Music , Stanford.

[11] Rajeswari Sridhar, S.R.Surya Dev, V. Sankar, K. Sivakumar. Automatic Tamil Lyric Generation based on neuro-linguistics. Proceedings of the International Conference on Informatics and Analytics. Article No. 129

[12] S. Rajeswari, D. J. Gladis, K. Ganga, and G. D. Prabha. Automatic tamil lyric generation based on ontological interpretation for semantics. Sadhana, 39(1):97–121, 2014.

[13] T. Padmapriya and V. Saminadan, “Inter-cell Load Balancing technique for multi-class traffic in MIMO-LTE-A Networks”, International Journal of Electrical, Electronics and Data Communication (IJEEDC), ISSN: 2320- 2084, vol.3, no.8, pp. 22-26, Aug 2015.

[14] S.V.Manikanthan and K.srividhya "An Android based secure access control using ARM and cloud computing", Published in: Electronics and Communication Systems (ICECS), 2015 2nd International Conference on 26-27 Feb. 2015,Publisher: IEEE,DOI: 10.1109/ECS.2015.7124833.