Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 3 (2023)

Leveraging Bi-Directional LSTM for Robust Lyrics Generation in Telugu: Methodology and Improvements

[1]VanamaYaswanth',[2]Ajay Kumar, [3]Neelesh Kumar Jain, [4]Nileshkumar Patel

[1,2,3,4] Computer Science, Jaypee University of Engineering and Technology, A.B.Road, Guna, 473226, M.P, India.

Abstract: The paper aims to analyze the various steps involved in creating semi-automated lyrics generators for Indian languages such as Telugu. Our study also examined the effects of bi-directional LSTM (long short-term memory) on different genres. After trying out several methods, we found that bi-directional LSTM works well with all formats. We improved our model with the help of 50,000 parameters during the training period and this is one of the largest crops for the Telugu language.

Keywords. *Lstm, Bi-Lstm, Nlp, Rnn, Generator*

1 Introduction

Artificial intelligence (AI) is now present in everything, including household appliances and phones. The most notable examples of AI applications are those found in autonomous vehicles, the healthcare sector, and telecommuting. In this paper, we ex- amine the use of Natural Language Processing (NLP) in the Telugu language. Humans use language as their primary tool to share their thoughts and emotions, and it differs from region to region. More than 7,000 languages are spoken worldwide, with the majority of them being spoken by small populations in the tropics. The nation with the widest variety of languages is Papua New Guinea [1].

Some well-known examples of conversational AI-based agents used by humans to connect with computers include Siri, Amazon Alexa, and Tesla [2, 3]. Music is used toimprove interpersonal relationships and individual experiences in some way [4]. Music is the generative link between humans and nature. There is no doubt about the calming effects of listening to music. It has been demonstrated that listening to music promotes mental relaxation [5].

Finding appropriate words for the song is the most important step if you want it to touch people's hearts and connect with the rest of the world. Most writer work extremely hard to provide additional lyrics in order to produce high-quality songs. In our application, we can see the solution to the issue that most lyricists and musicians are dealing with, which is how to generate new lyrics based on AI [14-15].

The remaining articles are organized as follows. Section 2 contains the detailed methodology used in this paper.Section 3 presents the overall network model used in the pro- posed framework for lyrics generation. Section 4 describes the implementation and testing of the proposed model. Finally, Section 5 and Section 6 provide concluding remarks and future directions for this work.

2 Methodology

2.1 System

The suggested model's workflow is depicted in Figure 1, which begins with data collection and progresses through pre-processing and building a deep learning model for creating new lyrics for the Telugu language.

2.2 Data Collection

Web scraping is the fundamental building block of these models. To collect data from many websites and web pages, we have utilized a variety of tools and advanced Python code using BS4[6]. We have extracted data from different sources and collected 2000 songs of different genres.

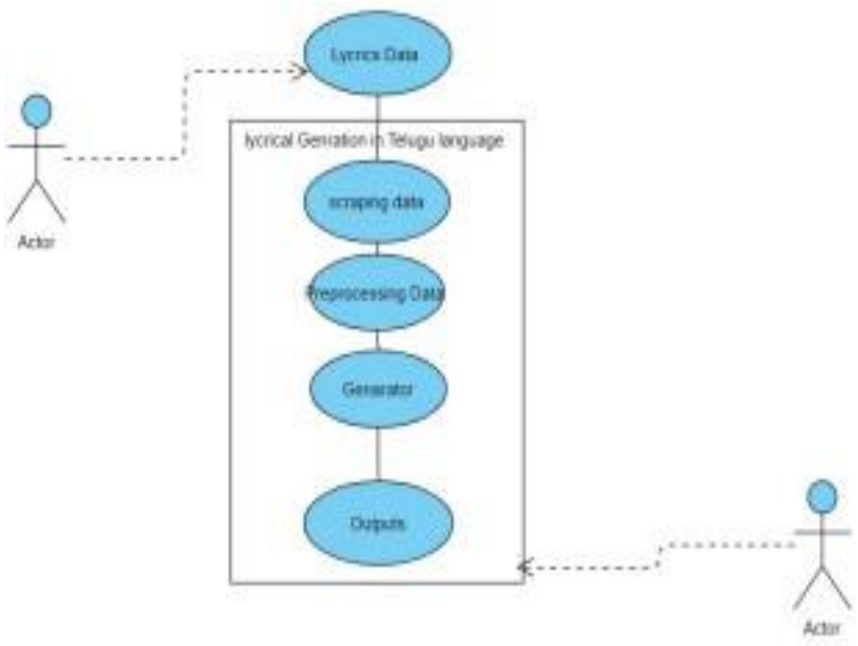

Figure 1. Use-Case Diagram.

2.3 Pre-processing

Pre-processing is fairly simple when dealing with the English language through NLTK or Spacy, two popular Python libraries for natural language processing [7]. While dealing with Indian languages, we used the Indic NLP-Library by IIIT Hyderabad to create our own tokenization and lemmatization for the Telugu language [8]. Regex was also used to eliminate unnecessary symbols from the text [9]. Results acquired after the preprocessing are 135 unique characters and 544217 corpus data as shown in Table1.

Table 1. Different Outcomes from the approach

Description	Outcomes
UniqueCharacter	135
Corpus	544217

2.3.1 Stop Words

While English, German, and French have stop words that are built into their grammar, Telugu has incredibly few stop words. This could change the meaning of a phrase if we eliminated several words in Telugu. As a result, we are very careful when dealing with stop words in Telugu.

2.3.2 Tokenization

The core of NLP is tokenization, and for the languages like English, German, and other languages, we use the NLTK Library. However, it was unknown for Indian languages until 2019, when Telugu, Hindi, and other Indian languages are tokenized. We have used Indic NLP-Library for the tokenization [8].

2.3.3 Normalization

A token's transformation into its base form is known as normalization. We eliminate the inflectional form of a word during normalization in order to achieve the base form by using the indic-NLP library.

Data Collection The first corpus is made up of the pre-processed data that was made from the data that was gathered through data scraping. This is one of the accomplishments achieved in this paper, when compared to the initial corpus produced by IIIT Hyderabad, which was around 3 lakhs, and the corpus developed for this project, which was 544436. A sample of data set of corpus is available at location[1].

3 Network Building

Over the past few years, there has been a significant advancement in the application of Bi-LSTM for text creation. Activation function and Dropmax, which are used in the first layer to reduce over-fitting, are not relevant in our case. In contrast, Soft max is useful when converting vectors in to probabilities. Because of the rare rhyme in our lyrics, these networks can- not recognize rhyme. In the next section, we'll examine various NLP approaches used to create text and why we used the Bi-LSTM approach.

3.1 Recurrent Neural Network (RNN):

The RNN is the basic neural network for processing sequential data. We also used it for forecasting data, with its limitations for small data sets only [10].

3.2 Long Short-Term Memory (LSTM):

The most popular learning concept in Deep learning is LSTM, which is very efficient for long sentences when it comes to Sequential Data. In LSTM, the only problem is that it can understand sentences as if they were paragraphs [10, 11].

3.3 Bi-Directional LSTM:

The BI-Directional Long Short-Term Memory (LSTM) model is a deep learning technique used to predict the next word in a sequence, such as a song. In the past, LSTM was used to train words in a single direction, but it proved to be less effective when we had long sentences [12,13]. Therefore, we have begun using Bi-Directional LSTM, which gives us the benefit of learning in both directions. Therefore, Telugu, Hindi, or any Indian language, makes more sense if we have a proper understanding of it. However, we need to learn it from both sides to be able to gain the benefits.

4 Testing implementation

The main objective of this manuscript is to obtain semi-automated lyrics, but after trying out several methods, we found that bi-directional LSTM works well with all formats. The BI-Directional LSTM model works by processing the input sequence in both forward and backward directions, allowing it to consider the context and dependencies of words both before and after the current word in the sequence. This enables the model to make more accurate predictions, as it can take into account the relationships between words in the sequence.To train the model, we used a Numpy array to find the shape and determine the number of trainable parameters. In our case, there were 595,587 trainable parameters out of a total of 595,587. The model was then built on different epochs, with good accuracy observed as the number of epochs in- creased.To predict the next word in a sequence, the BI- Directional LSTM model takes the input sequence and processes it through the forward and backward directions, using the trained weights and biases to make a prediction. This prediction is based on the relationships and dependencies between the words in the sequence, allowing the model to make an educated guess as to what the next word might be. As seen in, Figure 2, we have built our models on different epochs, and we have achieved good accuracy with respect ton epochs.

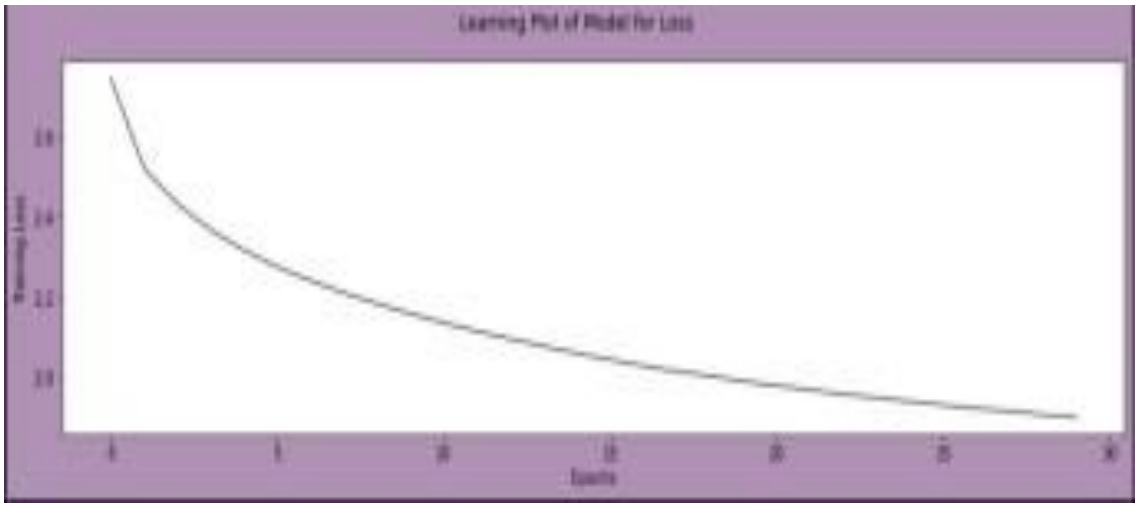

Figure 2. Loss function

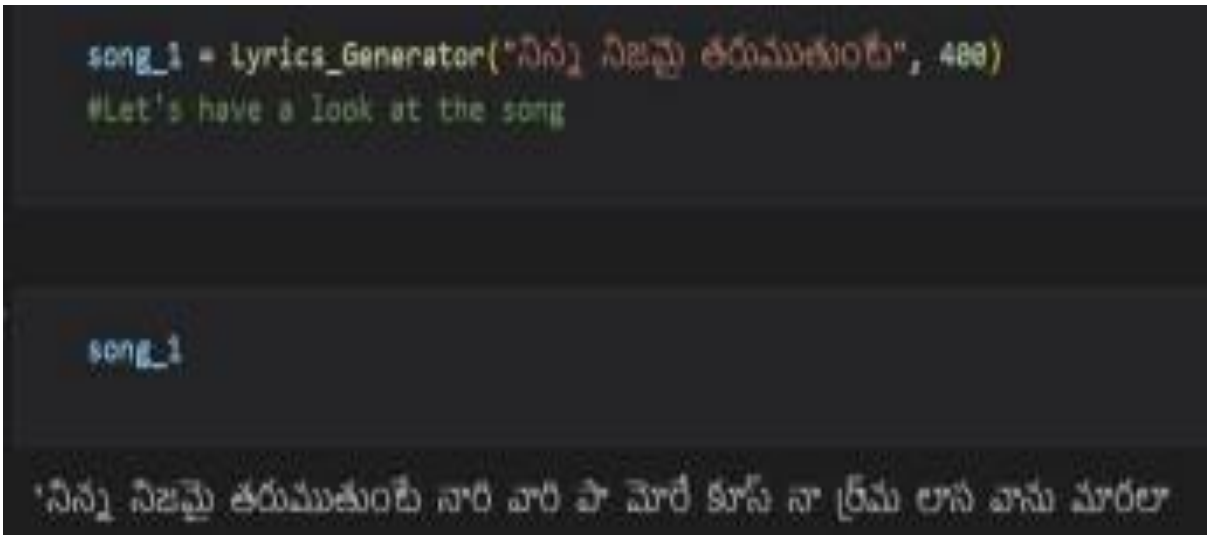

Figure 3. Output of the model

5 Conclusion:

Overall, the Bi-Directional LSTM model is a powerful tool for predicting the next word in a sequence, such as a song. By considering the context and dependencies of words both before and after the current word, it can make more accurate. We have created a semi-automated Telugu lyric's generator, as seen in Figure 3. It simply requires input, and you may specify the number of bytes you want to generate based on the number of characters in the lyrics; in this case, for instance, we've only used 400 characters.

6 Future Works:

We want to build text models using the LSTM as a first step, but after studying the GPT and GANS methods, we would prefer to increase text production using GANS and GPT-3 to represent other genres. Since our model is now based on genres, we would prefer to develop models that are based on specific genres in the future. It should incorporate every model. While we would also extend these models to other languages like Hindi, Tamil, and Kannada, this study primarily focuses on Telugu.

References

[1] Schreyer C and WagnerJ2022 Uncertainty in diversity: language shift and language planning in Papua New Guinea, a Kala case study. *Journal of Multilingual and Multicultural Development* 43(3): 262–275.

[2] Christian P. Janssen, Stella F. Donker, Duncan P. Brumby, Andrew L. Kun 2019 History and future of human-automation interaction, *International Journal of Human-Computer Studies*, Volume 131: 99-107.

[3] Brummelen J Van, Tabunshchyk V and Heng T 2021Alexa, Can I Program You?" Student Perceptions of Conversational Artificial Intelligence Before and After Programming Alexa". In: *Interaction Design and Children.* New York, NY, USA: association for Computing Machinery: 305–313.

DOI: 10.1145/3459990.3460730.

[4] Malloch S and Trevarthen C 2018 The Human Nature of Music. *Front. Psychol*. 9:1680. doi: 10.3389/fpsyg.2018.01680

[5] Blacking and Bruno 1995Nettl. *Music, Culture, and Experience: Selected Papers of John Blacking*. Ed. by and others. University of Chicago Press, 1995.

[6] Richardson L 2007.Beautifulsoup documentation.

URL: https://pypi.org/project/beautifulsoup4/.

[7] Spring R and Johnson M 2022 The possibility of improving automated calculation of measures of lexical richness for EFL writing: A comparison of the LCA, NLTK and SpaCy tools". *System* 106. DOI: 10.1016/j.system.2022.102770.

[8] Kakwani D. et al 2020 IndicNLPSuite: Monolingual corpora, evaluation benchmarks and pre-trained multilingual language models for Indian languages. In: *Findings of the Association for Computational Linguistics*: EMNLP 2020: 4948–4961.

[9] AhoA. V 1990 Algorithms for Finding Patterns in Strings, Handbook of Theoretical Computer Science Vol A". *A, ed. J. van Leeuwen, ElsevierSciencePublishers:* 257–297.

[10] Potash P, Romanov A, and Rumshisky A 2015 GhostWriter: Using an LSTM for Automatic Rap Lyric Generation. In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Lisbon, Portugal: Association for Computational Linguistics: 1919–1924. DOI: 10.18653/v1/D15-1221.

[11] Schmidhuber J, Gers F, and Douglas E 2002 Learning no regular languages: A comparison of simple recurrent networks and LSTM. *Neural computation* 14(9): 2039–2041.

[12] Dixit D, Bhagat A, and Dangi D 2022 Fake News Classification Using a Fuzzy Convolutional Recurrent Neural Network. *Computers, Materials and Continua 71:* 5733–5750. DOI: 10.32604/cmc.2022.023628.

[13] Ajay Kumar, Nileshkumar Patel, Neelesh Kumar Jain, Emotional Insights: Analysing Sentiments in Online Classes Amidst the COVID-19 Pandemic, TuijinJishu/Journal of Propulsion Technology Vol. 44 No. 3 (2023), pp. 1582-1592, https://doi.org/10.52783/tjjpt.v44.i3.544

[14] Ajay Kumar, Nileshkumar Patel, Neelesh Jain, The Impact of AI Technology on the Financial Sector and How AI is Modifying the State of Modern-Day Financial Institutions by Banking 4.0, Eur. Chem. Bull. 2023, 12 (Si6), 5448 – 5459.

[15] Nileshkumar Patel, Neelesh Jain, Ajay Kumar, Participation of AI Technologies in Financial Services DEFENSE AGAINST CYBER-ATTACKS, Eur. Chem. Bull. 2023, 12 (Si6), 5966 598.

Author Biography

B.Tech in 2023 from JUET Guna, Associate Manager @ iNeuron. His are of interest is Data Science and he is also interested in Helping the students with of Data science.

Ajay Kumar working as an Assistant Professor since 2006. He has completed his PhD in 2017 from Jaypee University of EngineeringandTechnologyinthedepartment of Computer Science and Engineering. His work area of Ph.D. was design and analysis of effective partitioned based clustering algorithm and its application. He has completed his M.E. from M.I.T.S. Gwalior in 2005. His M.E. thesis title was design and analysis of a data-mining tool. He has completed his B.Techin Information Technology from M.I.E.T., Meerut in 2002. His area of Interest includes data mining, image processing and patternrecognition.

Nilesh Kumar Patel has completed his B.E. (Computer Engineering) from Sardar Patel University, VV Nagar, Gujarat, M. Tech in Computer Science and Engineering from NIT, Bhopal and Ph.D. (Computer Science and Engineering) from Jaypee University of Engineering and Technology, Guna, MP.

His research and teaching proficiency area includes Internet of Things, Wireless Sensor Networks, Machine and Deep Learning, Systems and Network Programming, Distributed Systems, Database Systems, Computer Programming and Data Structures. Dr. Neelesh Kumar Jain did his Ph.D.in 2018 from Jaypee University of Engineering and Technology in the department of Computer Science and Engineering. The title of his research was "Efficient Approaches for Digital Image Forgery Detection". He received M. Tech. in Engineering Systems in 2006 from Dayal bagh Educational Institute, Agra and Bachelor's degree in Information Technology in 2003 from Dr. B.R.A. University, Agra.He has been recipient of Director'sGold Medal of Dayalbagh Educational Institute in M.Tech (2006). He has thirteen years of teaching experience for PG & UG courses of Computer Science & Engineering. He has published many research papers in reputed international journals and conferences including SCI indexed journals. His current research area is design an efficient algorithms, optimization techniques, Image forensics and machine learning.