

MULTILINGUAL STORY GENERATION AND SPEECH SYSTEM

¹CHITRADA ANKITH, ²K.RAJA RAJESWARI

¹Students, Department of MCA, B V Raju College, Bhimavaram Ap

²Assistant Professor, Department of MCA, B V Raju College, Bhimavaram Ap

ABSTRACT

With the rapid advancement of Artificial Intelligence and Natural Language Processing (NLP), automated content generation has become increasingly popular across various applications such as education, entertainment, and accessibility. This project presents a multilingual story generation and speech system that enables users to generate meaningful stories from simple keyword inputs and convert them into speech in multiple languages. The system supports languages such as English, Telugu, and Hindi, making it accessible to a diverse range of users. The proposed system utilizes transformer-based AI models combined with NLP techniques to process user-provided keywords. Initially, the input keywords are cleaned and structured using NLP preprocessing methods such as tokenization and normalization. The refined input is then passed to a transformer-based language model, which generates a coherent and contextually relevant story. Additionally, a text-to-speech (TTS) module is integrated to convert the generated text into audio output in the selected language, enhancing user interaction and accessibility. The system includes modules for user authentication, story generation, NLP processing, and speech

synthesis. Experimental results demonstrate that the system generates meaningful and grammatically correct stories in multiple languages with high efficiency. The speech output is clear and understandable, providing an engaging user experience. However, challenges such as maintaining contextual consistency and handling complex inputs remain. Overall, this project highlights the potential of AI and NLP in multilingual content generation and speech synthesis applications.

Keywords: *Natural Language Processing, Transformer Models, Multilingual Systems, Story Generation, Text-to-Speech, Artificial Intelligence, Language Models, Speech Synthesis.*

I.INTRODUCTION

In recent years, advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP) have significantly transformed the way humans interact with machines. Automated text generation systems are now widely used in applications such as chatbots, content creation, virtual assistants, and educational tools. One such emerging application is story generation, where AI models can generate meaningful and creative

stories based on user input. However, most existing systems are limited to a single language, which restricts their accessibility and usability for a diverse population. Additionally, integrating speech capabilities into such systems can further enhance user experience, especially for users with reading difficulties or language barriers.

The proposed system focuses on developing a multilingual story generation and speech system that allows users to generate stories in multiple languages such as English, Telugu, and Hindi. The system uses transformer-based AI models, which are highly effective in understanding and generating natural language. Initially, user-provided keywords are processed using NLP techniques such as tokenization, normalization, and stop-word removal. These processed inputs are then fed into the transformer model to generate a coherent and contextually relevant story. The system also integrates a text-to-speech (TTS) module, which converts the generated text into audio output in the selected language, making the system more interactive and accessible.

The system is designed with multiple modules, including user login, story generation, NLP processing, and speech synthesis. Users can input keywords, select their preferred language, and choose whether to enable speech output. Although the system performs effectively, challenges such as maintaining story coherence

for complex inputs and handling diverse linguistic structures remain. Future enhancements may include support for more languages, improved contextual understanding, and real-time voice interaction. Overall, this project demonstrates the potential of combining AI and NLP techniques to create intelligent, multilingual, and user-friendly content generation systems.

II SURVEY OF RESEARCH

The study by A. Vaswani et al. (2017) [1] introduced the Transformer architecture, which revolutionized natural language processing tasks. The methodology uses attention mechanisms to process sequential data more efficiently than traditional recurrent models. Results showed significant improvements in text generation and language understanding tasks. However, transformers require large datasets and computational resources. This research forms the foundation for story generation in the proposed system.

The work by T. Mikolov et al. (2013) [2] introduced word embedding techniques such as Word2Vec. The methodology converts words into vector representations that capture semantic relationships. Results demonstrated improved performance in various NLP tasks. However, embeddings may not fully capture contextual meaning. This study supports keyword processing and feature representation in the system.

The study by J. Devlin et al. (2018) [3] proposed BERT (Bidirectional Encoder Representations from Transformers) for improved language understanding. The methodology uses bidirectional training to capture context from both directions in a sentence. Results showed state-of-the-art performance in NLP tasks. However, high computational cost is a limitation. This research supports advanced text generation capabilities.

The research by I. Goodfellow et al. (2016) [4] discussed deep learning techniques for natural language processing. The methodology uses neural networks to learn hierarchical representations of language. Results demonstrated improved accuracy and efficiency in text-related tasks. However, large datasets are required. This study provides the theoretical basis for AI-based story generation.

The study by D. Jurafsky and J. H. Martin (2009) [5] introduced fundamental concepts of speech processing and NLP. The methodology includes techniques for text analysis and speech synthesis. Results showed effective integration of NLP and speech systems. However, speech quality depends on model design. This research supports the text-to-speech module in the system.

The work by Y. Wang et al. (2017) [6] explored neural text-to-speech systems. The methodology uses deep learning to convert text

into natural-sounding speech. Results demonstrated improved speech quality and naturalness compared to traditional methods. However, computational complexity is high. This study supports the speech generation component of the proposed system.

III. WORKING METHODOLOGY

The proposed multilingual story generation and speech system follows a structured methodology to convert user-provided keywords into meaningful stories and audio output. Initially, the process begins with user input and preprocessing. The user enters keywords and selects the desired language (English, Telugu, or Hindi) along with the speech option. The input keywords are processed using Natural Language Processing (NLP) techniques such as tokenization, stop-word removal, and normalization. This step ensures that irrelevant or noisy words are removed and the input is transformed into a clean and structured format suitable for further processing. Proper preprocessing is essential for improving the quality and relevance of the generated story.

In the next phase, the processed keywords are fed into a transformer-based AI model for story generation. The transformer model uses attention mechanisms to understand the context and relationships between words, enabling it to generate coherent and meaningful stories. The model is trained on large text datasets, allowing

it to produce grammatically correct and contextually relevant outputs. The generated story is then formatted and displayed to the user in the selected language. This phase is crucial as it ensures that the story aligns with the input keywords and maintains logical flow and creativity.

Finally, the system integrates a text-to-speech (TTS) module to convert the generated story into audio output. If the user selects the speech option, the system processes the text using a speech synthesis model to generate voice output in the chosen language. The audio is then played through the application interface, enhancing user interaction and accessibility. The system also allows users to generate multiple stories with different inputs and languages. Despite its effectiveness, challenges such as maintaining long story coherence and handling complex multilingual structures remain. Future improvements may include real-time voice input, enhanced language support, and more advanced transformer models for better performance.

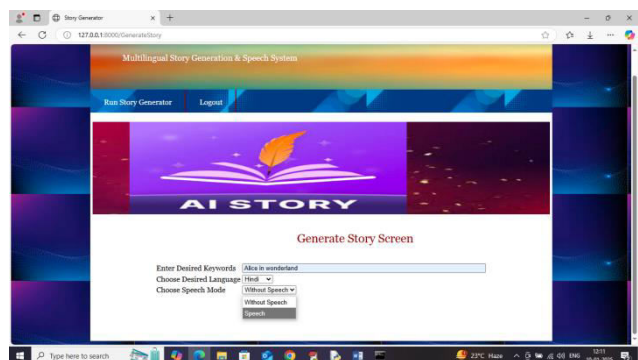
IV RESULTS EXPLANATIONS

In propose work we are designing AI and NLP based story generator where user can input desired keywords and then AI transformer algorithm will apply NLP (natural language API) to clean keywords and then generate precise story based on given keywords.

User can generate story in English, Telugu or Hindi Language and can make select to speech generated story in selected language.

To implement this project we have designed following modules

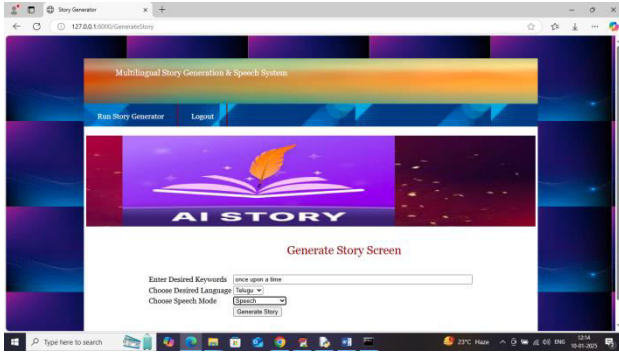
- 1) User Login: user can login to system
- 2) Generate Story: using this module user can enter desired keywords and then select desired language and speech mode to generate and speech story in desired language.
- 3) NLP & AI algorithm: given keywords will be input to Transformer AI algorithm to clean keywords and then generate precise story based on keywords.



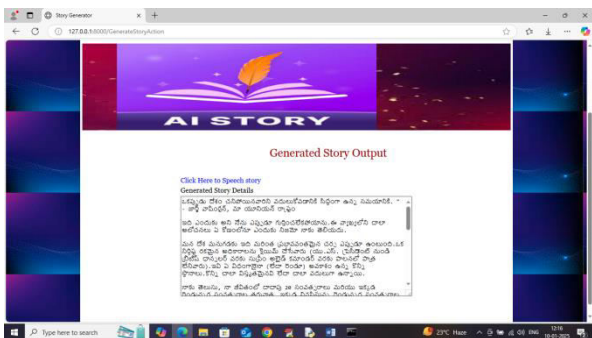
In above screen I selected language as 'Hindi' along with speech mode and then press button to get below page



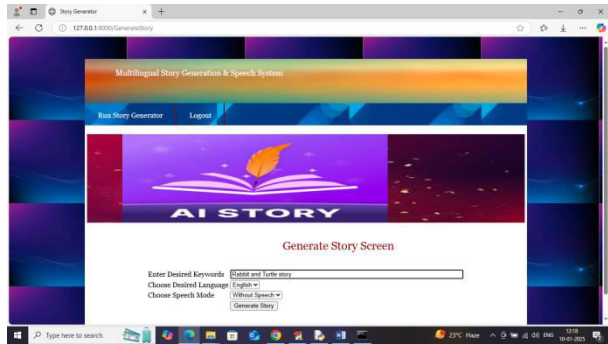
In above screen in text area can see generated story and can click on blue text ‘Click Here to speech story’ to start playing story and now try another language



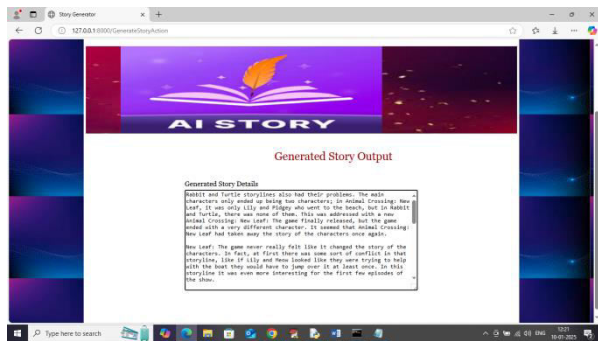
In above screen I gave some other keywords along with speech and Telugu language and below is the output



In above screen in text area can see story generated in Telugu and can click on ‘Blue colour text’ to play story and now click generate story in English



In above screen asking to generate story on ‘Rabbit and Turtle’ in English language without speech and below is the output



In above screen story generated on rabbit and turtle. Similarly enter any keywords and then generate story in desired language by following above screens

V.CONCLUSION

The proposed multilingual story generation and speech system successfully demonstrates the integration of Artificial Intelligence and Natural Language Processing to generate meaningful stories and convert them into speech across multiple languages. By utilizing transformer-based models, the system effectively processes user-provided keywords and produces contextually relevant and

grammatically correct stories in English, Telugu, and Hindi. The addition of a text-to-speech module enhances user interaction by enabling audio playback, making the system more accessible and engaging for a wide range of users, including children and visually impaired individuals.

The experimental results indicate that the system performs efficiently in generating diverse and creative stories while maintaining linguistic accuracy. The NLP preprocessing stage improves the quality of input, leading to better story generation. The speech output is clear and understandable, further enhancing the user experience. However, the system has certain limitations, such as variations in story coherence for complex inputs and differences in speech quality across languages.

Future work can focus on improving contextual understanding, supporting additional languages, and integrating real-time voice input for a more interactive experience. Advanced transformer models and improved speech synthesis techniques can further enhance system performance. Overall, this project highlights the potential of AI-driven multilingual systems in content generation and accessibility, providing a scalable and user-friendly solution for storytelling applications.

RE.FERENCES

- [1] A. Vaswani et al., “Attention is all you need,” in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2017, pp. 5998–6008.
- [2] T. Mikolov et al., “Efficient estimation of word representations in vector space,” in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2013.
- [3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.
- [4] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [5] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 2nd ed. Upper Saddle River, NJ, USA: Pearson, 2009.
- [6] Y. Wang et al., “Tacotron: Towards end-to-end speech synthesis,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2017, pp. 4006–4010.
- [7] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann, 2011.
- [8] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.

- [9] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2010.
- [10] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.
- [11] J. Leskovec, A. Rajaraman, and J. Ullman, *Mining of Massive Datasets*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
- [12] E. Alpaydin, *Introduction to Machine Learning*, 4th ed. Cambridge, MA, USA: MIT Press, 2020.
- [13] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. New York, NY, USA: Springer, 2009.
- [14] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*. New York, NY, USA: Springer, 2013.
- [15] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 1995, pp. 1137–1143.
- [16] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.
- [17] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [18] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [19] F. Provost and T. Fawcett, *Data Science for Business*. Sebastopol, CA, USA: O'Reilly Media, 2013.
- [20] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
- [21] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
- [22] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. Sebastopol, CA, USA: O'Reilly Media, 2017.
- [23] F. Chollet, *Deep Learning with Python*. Shelter Island, NY, USA: Manning Publications, 2017.
- [24] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [25] A. Graves, "Generating sequences with recurrent neural networks," arXiv:1308.0850, 2013.