AI – POWERD MEME GENARATOR

Amar H. M.

Department of AI & ML

Dayananda Sagar University

Bengaluru India

eng23am1001@dsu.edu.in

Sanjay S.

Department of AI & ML

Dayananda Sagar University

Bengaluru India

eng23am1006@dsu.edu.in

Deepak Mahto.

Department of AI & ML

Dayananda Sagar University

Bengaluru India

eng22am0087@dsu.edu.in

Index Terms

LSTM, Meme Generation, Deep Learning, Natural Language Processing, Streamlit, Text Generation, Neural Networks, Captioning Model, Machine Learning, Image-Based Content Creation.

Abstract

Memes have evolved into a powerful medium of expression across digital platforms, blending humor with social commentary. Creating meaningful memes, however, typically requires both creativity and time. This paper presents an AI-powered meme generation system that automates the caption creation process using a Long Short-Term Memory (LSTM) network combined with a fallback template-based captioning mechanism. The LSTM model extends user-provided seed text to generate context-aware captions, while a rule-based filter ensures linguistic correctness by removing repeated, irrelevant, or non-English outputs.

When the generated text does not meet quality criteria, the system employs predefined, professional English caption templates to maintain reliability. The complete solution is deployed through a Streamlit interface enabling users to upload images or select preset templates, enter topics, and instantly generate memes. Experimental results demonstrate that this hybrid approach provides fast, readable, and visually appealing meme outputs while maintaining system simplicity and low computational requirements. The proposed method offers a practical and efficient tool for automated meme creation and lays the foundation for future upgrades involving transformer-based models and image-aware captioning.

I. Introduction

Memes have become a central component of modern digital communication, serving as a creative and humorous way to express ideas, emotions, and social perspectives. Their widespread presence across social media platforms highlights the growing demand for tools that simplify and enhance meme creation. Traditionally, generating a meme involves manually selecting an image, crafting an appropriate caption, and formatting the content visually. This process requires creativity, linguistic clarity, and time, which may not always be accessible to every user.

With advances in artificial intelligence and natural language processing, automated text generation has become increasingly effective. Among various sequence models, Long Short-Term Memory (LSTM) networks have proven capable of understanding contextual relationships within text, enabling them to generate coherent word predictions. Leveraging this capability for meme generation offers a promising approach to reduce user effort while preserving creativity.

This paper presents an AI-powered meme generator that combines an LSTM-based caption extension model with a rule-based quality filter and a template-driven fallback captioning approach. The hybrid design ensures that the system produces readable, humorous, and context-appropriate captions, even when the model output is insufficient. The final implementation is deployed through a Streamlit web interface, providing users with an interactive environment to upload images, choose templates, input topics, and generate memes instantly.

The proposed system demonstrates how lightweight deep learning models, when paired with robust filtering and interface design, can effectively support creative digital tasks. This work highlights the potential of integrating AI-driven text generation into everyday content creation tools, providing users with an accessible and efficient meme production experience.

II. Literature Review

Early work in text generation relied heavily on statistical language models such as n-grams, which model word sequences based on fixed-length histories. Although computationally simple, these models suffered from sparse data issues and lacked the ability to capture long-range dependencies within a sentence. As a result, they produced repetitive or contextually weak text, limiting their usefulness in creative applications such as caption or meme generation.

The introduction of Recurrent Neural Networks (RNNs) marked a significant improvement in sequence modeling by enabling dynamic context retention. However, standard RNNs were prone to vanishing and exploding gradients, which restricted their ability to learn long-term patterns. This challenge led to the development of Long Short-Term Memory (LSTM) networks, which introduced gated memory units capable of maintaining information across extended time steps. Research has shown that LSTMs perform well on tasks such as language modeling, sentiment analysis, and caption generation due to their improved contextual understanding.

In the area of visual-text synthesis, studies have explored pairing images with textual descriptions using encoder–decoder architectures and attention mechanisms. While these models achieve high-quality results, they typically require large datasets and substantial computational resources. Emerging transformer-based architectures such as GPT, BERT, and ViT have further advanced generative capabilities, but their complexity and resource requirements make them less suitable for lightweight or rapid prototyping applications.

Work on meme automation specifically has investigated humor generation, rule-based caption creation, and template recommendation systems. Many systems rely strictly on template libraries or handcrafted rules, which limits diversity and creativity. Hybrid systems combining AI models with controlled templates have been shown to improve consistency while maintaining output quality.

Building on these developments, the present work adopts a balanced approach by using an LSTM model for creative caption extension and integrating it with rule-based verification and professional fallback templates. This helps ensure that the system remains computationally efficient while still delivering coherent, readable, and relevant meme captions.

III. Problem Statement

Meme creation is a highly popular form of digital expression, yet the process typically requires users to manually craft captions that are humorous, contextually relevant, and linguistically coherent. Many individuals lack the time, creative inspiration, or language skills needed to produce high-quality memes consistently. Existing automated meme generators often depend solely on predefined templates or rule-based caption libraries, resulting in repetitive and predictable outputs. Conversely, systems that rely entirely on deep learning models may produce grammatically incorrect, irrelevant, or non-English captions, especially when trained on limited datasets.

Therefore, there is a clear need for a hybrid meme-generation system that can:

1. generate meaningful, English captions automatically based on user-provided topics,
2. ensure linguistic correctness and contextual relevance,
3. avoid repetition or nonsensical outputs produced by raw model predictions, and
4. operate efficiently on lightweight infrastructure without requiring large-scale transformer models.

The problem addressed in this work is to design and implement an AI-driven meme generator that balances creativity and reliability by combining an LSTM-based caption extension model with rule-based quality checks and template-based fallback mechanisms. This aims to provide users with a consistent, accessible, and easy-to-use tool for automated meme creation.

IV. Objectives

The primary objective of this work is to develop an AI-powered meme generation system that automates caption creation while ensuring clarity, relevance, and usability. To achieve this goal, the system is designed with the following specific objectives:

1. **Automate Caption Generation:** Develop an LSTM-based language model capable of extending user-provided seed text into meaningful, context-aware captions suitable for meme creation.
2. **Ensure Caption Quality and Reliability:** Implement a rule-based filtering mechanism to eliminate non-English, repetitive, irrelevant, or low-value words produced by the model.
3. **Provide a Consistent Fallback System:** Integrate a professional template-based caption generator to guarantee readable and coherent output when the LSTM model fails to produce acceptable text.
4. **Enable Flexible Image Selection:** Allow users to either upload custom images or choose from existing templates to suit different meme styles and use cases.
5. **Deliver a User-Friendly Interface:** Build an interactive Streamlit-based application that simplifies meme creation through intuitive controls, previews, and seamless output generation.
6. **Optimize for Lightweight Deployment:** Design the system to operate efficiently without the need for large transformer models, making it accessible for rapid prototyping and low-resource environments.
7. **Support Real-Time Meme Creation:** Ensure that caption generation and image rendering occur quickly enough for practical, real-world usage.

V. Dataset Description

The effectiveness of the proposed meme caption generator depends heavily on the quality and structure of the textual dataset used to train the LSTM model. Since memes typically rely on short, expressive, and conversational language, the dataset was prepared to capture these characteristics while maintaining clarity and simplicity.

The dataset consists of **text-based meme captions**, **short humorous statements**, and **topic-oriented phrases** collected from openly available meme text sources and curated caption lists. All gathered text was cleaned, standardized, and formatted into plain sentences suitable for sequence modeling. Non-English content, emojis, long paragraphs, and irrelevant symbols were removed to ensure consistency during training.

To improve model performance, the dataset was further preprocessed by converting sentences to lowercase, tokenizing words, and removing extremely rare or overly common tokens that do not contribute meaningfully to caption structure. The final processed dataset was used to create a tokenizer with a fixed vocabulary size and compute the maximum sequence length required for padding.

Since meme text is typically short, the dataset focuses on **brief phrases ranging from 5 to 15 words**, allowing the LSTM to learn natural transitions between words without overfitting on long-form text. This makes the model lightweight and well-suited for rapid caption extension based on user-provided seed text.

Overall, the dataset was designed to balance humor, simplicity, and contextual relevance, enabling the LSTM model to generate captions that align with the style of modern meme culture while maintaining grammatical correctness.

VI. Proposed Methodology

The proposed system follows a structured, multi-stage process designed to generate high-quality memes using deep learning. Each step builds on established practices while integrating modern techniques to ensure consistency, creativity, and reliability.

1. Data Collection

A curated meme dataset is gathered, containing:

- Popular meme templates
- Existing meme images
- Associated captions
  This provides the foundation for understanding language patterns and visual context.

2. Data Pre-Processing

Before training, the dataset is refined through several steps:

- **Text Cleaning:** Removing special characters, unnecessary symbols, and formatting noise.
- **Tokenization:** Breaking the caption text into sequences suitable for model training.
- **Padding:** Ensuring all sequences match the required length for LSTM input.
- **Image Normalization:** Resizing and scaling template images for uniform processing.

3. Feature Extraction

Image features are extracted using traditional image preprocessing and, if required, a lightweight CNN.
Text features are converted into numerical embeddings using word embeddings or token IDs.

4. LSTM-Based Caption Generation

A Long Short-Term Memory (LSTM) model serves as the core engine for text generation.

- It learns language structure, humor patterns, and caption styles from the dataset.
- It predicts the next word in a sequence, forming complete captions.
- This helps maintain contextual flow and coherence in the generated text.

5. Meme Template Selection

Based on either user input or model interpretation, the system identifies the most suitable meme template.
Templates are selected considering:

- Caption sentiment
- Category
- Visual style
  This ensures harmony between the caption and the image.

6. Caption Rendering on Image

The generated caption is placed on the chosen meme template. This involves:

- Text rendering
- Font selection
- Positioning the caption in standard meme format
  The output must resemble commonly accepted meme styles.

7. Streamlit-Based Deployment

The complete system is deployed using Streamlit for a simple, functional interface:

- User submits text or triggers random caption generation
- Back-end model processes data and prepares the meme
- The final output is displayed for download

8. Evaluation & Testing

To maintain output quality, the system is assessed based on:

- Caption relevance
- Readability
- User satisfaction

This ensures that the model behaves consistently.

VII. Conclusion

This project demonstrates a practical and creative application of deep learning by developing an AI-powered meme generator using LSTM and Streamlit. Through careful dataset preparation, text preprocessing, sequence modeling, and template rendering, the system is able to produce meaningful and contextually relevant meme captions. By combining traditional sequence-based learning with a simple and accessible deployment platform, the model offers a balanced blend of innovation and usability.

The work highlights how classical neural network architectures such as LSTM can still perform effectively for structured text generation tasks. The system's straightforward workflow—ranging from data preparation to deployment—ensures that it remains easy to understand, expand, and maintain. While the results are promising, further improvements such as integrating transformer-based models or expanding the dataset can strengthen caption quality and humor consistency.

Overall, this project serves as a strong step toward automated content generation and showcases how established machine learning techniques can be applied in a modern, user-friendly setting.

VIII. Future Work

- Implement deep learning models (LSTM, BERT, transformers) for improved accuracy.
- Expand the dataset using larger and diverse real-world email samples.
- Add URL reputation checks and attachment scanning for deeper threat analysis.
- Integrate authentication protocols such as SPF, DKIM, and DMARC.
- Deploy as a browser extension or integrate with enterprise email servers.
- Add advanced MLOps features like drift detection and automated model versionin

IX. ACKNOWLEDGMENT

The authors express their sincere gratitude to **Dr. Bahubali** for his valuable guidance, encouragement, and insightful feedback throughout the development of this work. The authors also thank the Department

of Artificial Intelligence & Machine Learning, Dayananda Sagar University, for providing the necessary resources and support. Their continued motivation and technical assistance played a vital role in the successful completion of this project

X. REFERENCES

[1] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, pp. 436–444, 2015.

[3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv:1409.1556, 2014.

[4] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," arXiv:1301.3781, 2013.

[5] F. Chollet, *Deep Learning with Python*, 2nd ed., Manning Publications, 2021.

[6] Streamlit Inc., "Streamlit Documentation," Available: https://docs.streamlit.io/

[7] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[8] D. Jurafsky and J. Martin, *Speech and Language Processing*, 3rd ed. (Draft), 2023.

[9] A. Karpathy, "The Unreasonable Effectiveness of Recurrent Neural Networks," 2015. (Blog)

[10] J. Brownlee, "Deep Learning for Natural Language Processing," Machine Learning Mastery, 2018