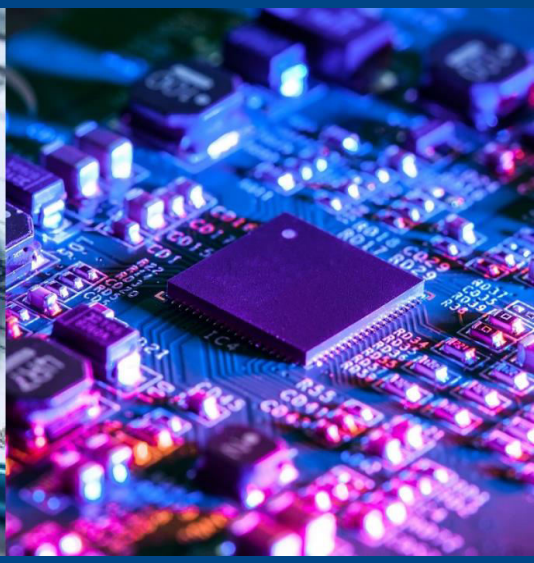
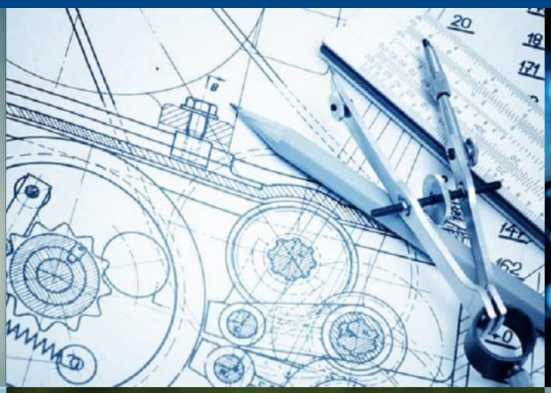
ISSN: 2582-7219

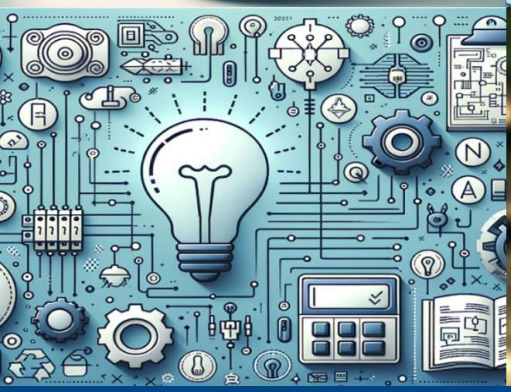




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206 **Volume 8, Issue 5, May 2025**

© 2025 IJMRSET | Volume 8, Issue 5, May 2025| DOI:10.15680/IJMRSET.2025.0805121

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Intelligent Writing Assistant Using NLP & Transformation – Based Grammar Correction

V Hemalatha, K Rukkumani

Assistant Professor, Dept. of CSE, NSNCET, Karur, India

Student, Dept. of CSE, NSNCET, Karur, India

ABSTRACT: In today's digital world, effective English communication is essential for education, business, and daily interactions. This project aims to develop a web-based English Text Correction System that enhances user-submitted content by correcting grammar, spelling, punctuation, and sentence structure while providing meaningful suggestions. Using Natural Language Processing (NLP) techniques and advanced grammar-checking algorithms, the system will analyze input text, detect errors, and generate a refined version with improved readability and coherence. Users can input their content into a text box and click a correction button to receive suggestions or an enhanced version of their text.

Key features include grammar and spelling correction, appropriate word usage, sentence structure improvements, and verb tense adjustments. The system will enhance punctuation for readability and optimize sentence structure for clarity and coherence. An optional error-highlighting feature will help users visualize mistakes, while educational explanations for corrections will offer insights into grammar rules and sentence restructuring. Designed as a mastering-oriented device, it'll not only correct errors but also assist improve customers' writing capabilities. With a user-friendly interface, the platform will offer actual-time corrections, causes, and hints—making it an reachable and interactive solution for students, experts, and writers in search of to refine their English talent.

KEYWORDS: Natural Language Processing (NLP), Text Correction, Grammar Checking, Spelling Correction, AI Writing Assistant.

I. INTRODUCTION

In today's digital era, effective written communication is essential across academic, professional, and personal domains. However, many users struggle with grammar, sentence structure, and clarity when writing in English. To address this challenge, the proposed project introduces an intelligent grammar correction system powered by Natural Language Processing (NLP) and deep learning. This system utilizes a pretrained T5 transformer model to identify and correct grammatical errors in user-submitted text. Built with Streamlit, the application offers a simple, user-friendly interface that provides instant grammar correction without requiring complex setups or user authentication. The goal is to deliver a practical, lightweight solution that enhances writing quality and supports users in producing grammatically accurate content.

II. LITERATURE REVIEW

In recent years, Natural Language Processing (NLP) and artificial intelligence (AI) have seen remarkable progress in the field of grammar correction and writing assistance. Several commercial and academic tools have been developed to identify and correct grammar, punctuation, spelling, and sentence-level issues. Traditional tools like Grammarly and LanguageTool are widely used, leveraging rule-based approaches and shallow machine learning techniques to provide basic corrections. However, these systems often lack contextual awareness and struggle with complex or nuanced grammar issues. With the advent of transformer-based models like BERT, RoBERTa, and T5, researchers have shifted toward deep learning methods that understand sentence structure holistically, enabling more accurate and context-aware grammar correction.

Raffel et al. proposed the Text-to-Text Transfer Transformer (T5) [1], which unified various NLP tasks into a single framework. T5 treats grammar correction as a sequence-to-sequence task where input text is transformed into a corrected output. This foundational concept inspired many grammar correction models, including the one used in this project.

Another key model is BERT by Devlin et al. [2], which uses bidirectional context to improve understanding in sentence-level tasks. While BERT is primarily used for classification tasks, its encoder structure has influenced hybrid grammar correction frameworks that combine detection and correction modules.

Junczys-Dowmunt et al. [3] proposed a neural-based approach using Transformer models for Grammatical Error Correction (GEC). Their work showed that even without rule-based systems, neural models trained on large GEC datasets like CoNLL-2014 and JFLEG can outperform traditional tools by understanding syntactic and semantic patterns.

Grammarly [4], one of the most popular commercial tools, uses a combination of deep learning, statistical models, and rule-based grammar checks. While effective for general use, its inner workings are proprietary, and it often lacks transparency and openness for customization in academic or developer environments.

Language Tool [5] is another widely used open-source grammar correction tool that applies pattern matching and rules. Although it performs well on standard errors, it struggles with deep semantic understanding and tends to overcorrect or miss contextually complex issues.

Relevance to current Research

This project builds on the strengths of transformer-based models by implementing a pretrained T5 model from Hugging Face to perform real-time grammar correction. Unlike Grammarly or LanguageTool, which are either closed-source or rule-based, the current system leverages deep contextual understanding provided by T5.

By using open-source libraries such as PyTorch and Hugging Face Transformers, this project aligns with current research trends while offering flexibility for future enhancements like multilingual support, sentence-level explanations, and improved error identification. It bridges the gap between academic research and practical usability by deploying the model through a lightweight and accessible Streamlit interface.

Summary Table: Related Work

No.	Paper / Tool	Authors	Key Points	Relevance to Current Work
1	Text-to-Text Transfer Transformer (T5)	Raffel et al., 2020	Unified NLP model that treats grammar correction as a sequence-to-sequence task	Forms the core model used in the system
2	BERT: Bidirectional Transformers	Devlin et al., 2019	Improved context understanding through bidirectional attention	Influenced sentence representation techniques
3	Neural GEC Using Transformers	Junczys-Dowmunt et al., 2018	Demonstrated end-to-end grammar correction using neural methods	Validates use of pretrained transformers for GEC
4	Grammarly	Grammarly Inc.	Commercial tool using deep learning + rules	Sets a functional benchmark but lacks customization
5	LanguageTool	Open-source project	Rule-based grammar correction engine	Highlights limitations of static rule-based systems

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. METHODOLOGY OF PROPOSED SURVEY

The proposed model introduces an AI-driven grammar correction system that leverages the capabilities of a pretrained T5 transformer model to provide accurate and context-aware text corrections. Unlike traditional rule-based tools, this system processes the entire sentence to understand context, grammar patterns, and structure before generating corrections. Developed using Python and deployed through Streamlit, the platform offers a real-time, interactive interface where users can input text and instantly receive corrected output. The model operates locally and does not require server-side APIs or databases, making it lightweight and accessible. Although adaptive learning is not yet implemented, the system is designed for future scalability, where user feedback could be integrated to improve correction accuracy over time. This approach offers a practical and efficient solution for students, professionals, and anyone seeking to enhance their English writing skills.

IV. SYSTEM ARCHITECTURE

The system follows a structured and sequential architecture to process and correct text efficiently using modern NLP techniques. The user begins by entering a sentence or paragraph into the Streamlit web interface. Upon clicking the correction button, the input text is preprocessed by appending a task prefix ("gec:") and tokenizing it using the Hugging Face T5Tokenizer. This tokenized input is passed directly to a pretrained T5 grammar correction model hosted locally using PyTorch. The model processes the text and returns a grammatically corrected version based on its learned language patterns. To identify specific mistakes, the corrected output is compared with the original input using Python's difflib library. This highlights word-level differences, which are then displayed alongside the corrected text. The entire interaction happens within the Streamlit app, providing an end-to-end grammar correction experience in real time without the need for any external API or database.

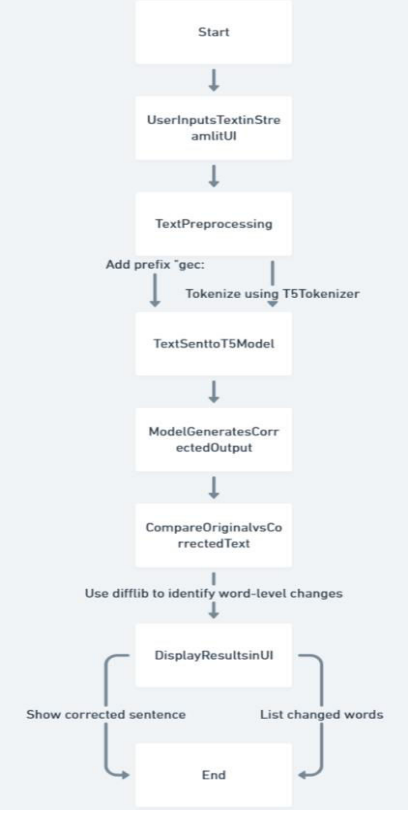


Fig 1 Work Flow

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

MODULES

Your grammar correction system can be broken down into the following core modules. Each module plays a specific role in handling input, processing it with the AI model, and displaying the output to the user.

1 Text Input Modules

This module is responsible for capturing user input through the Streamlit interface. It presents a text area where users can type or paste the sentence or paragraph they want to check. It also handles the submission event when the user clicks the "Correct Grammar" button. The input collected here becomes the foundation for the subsequent processing steps.

2 Pre – Processing Modules

Once the user submits the text, this module prepares the data for the AI model. It adds a task-specific prefix "gec:" to the input text to indicate a grammar correction task and then tokenizes the sentence using Hugging Face's T5Tokenizer. This ensures that the input is in the correct format for the transformer model to understand and process.

3 Grammar Correction Modules

This is the core module where the actual grammar correction takes place. The tokenized input is passed to the pretrained T5 model (prithivida/grammar_error_correcter_v1), which is responsible for generating a corrected version of the sentence. The model is loaded and executed using PyTorch, and it returns the refined output after analyzing the grammatical structure and context of the input sentence.

4 Error Identification Modules

After receiving the corrected output, this module compares the original and corrected sentences to identify any differences. It uses Python's difflib library to perform a word-level comparison and generate a list of words that were changed during the correction. This helps in highlighting the specific grammar errors without needing deep parsing or rule explanations.

5 Result Display Modules

This module handles the presentation of results to the user through Streamlit. It displays the corrected version of the text using success or info boxes. If any mistakes were detected, they are shown using warning boxes. This module ensures that the output is clear, readable, and visually distinct, helping users quickly understand what changes were made.

V. SYSTEM IMPLEMENTATION

The following screenshots demonstrate the functioning of the Intelligent Writing Assistant system. The user inputs text into the interface, and the system provides grammar, spelling, and sentence structure corrections in real time.

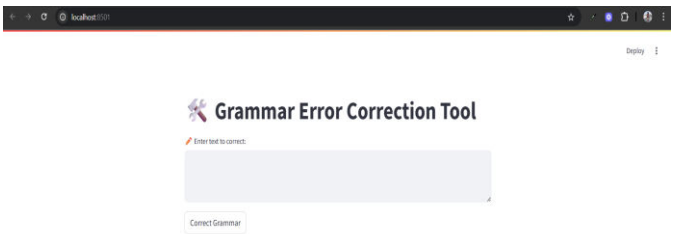


Fig 2 Streamlit Interface - Text Input Screen

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

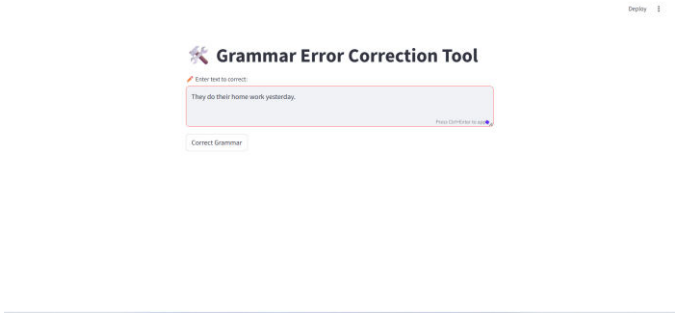


Fig 3 User Input

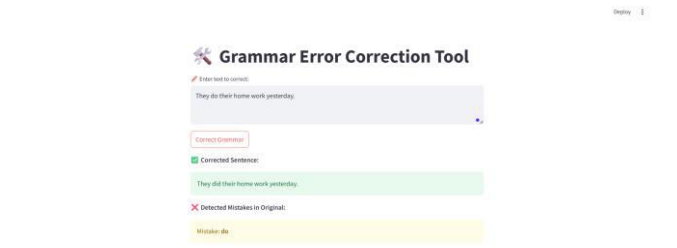


Fig 4 Corrected Output Display with Suggestions

VI. CONCLUSION AND FUTURE WORK

The Intelligent Grammar Correction System effectively applies advanced NLP techniques using a pretrained T5 transformer model to deliver real-time, context-aware grammar corrections through a simple Streamlit interface. Its open-source, lightweight design makes it accessible to users without technical setup. Unlike traditional rule-based tools, it understands sentence context for more accurate results. While currently limited to English and basic correction display, the system offers strong potential for future enhancements such as multilingual support, detailed error explanations, and adaptive learning, positioning it as a practical and scalable AI-powered writing assistant.

REFERENCES

1. Vaswani, N. Shazeer, N. Parmar, et al., “Attention is All You Need,” in Proc. Advances in Neural Information Processing Systems (NeurIPS), pp. 5998–6008, 2017.
2. J. Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in Proc. Conf. North American Chapter of the Association for Computational Linguistics (NAACL-HLT), pp. 4171–4186, 2019.
3. Raffel, N. Shazeer, A. Roberts, et al., “Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer,” J. Machine Learning Research (JMLR), vol. 21, no. 140, pp. 1–67, 2020.
4. T. Brown, B. Mann, N. Ryder, et al., “Language Models are Few-Shot Learners,” in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 1877–1901, 2020.
5. S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com