AUTO CORRECT TOOL

[1]Momin Mohammed Rehan, [2]Umer Khan
Mr. Suhas Salve
MGM'S College of Engineering, Nanded
[1]abdulkabir0717@gmail.com, [2]s21_umer_khan@mgmcen.ac.in

Abstract
The Autocorrect Tool is an intelligent text correction system designed to enhance user writing by automatically identifying and correcting spelling and grammatical errors in real-time. This system addresses the common challenges faced by users in maintaining high-quality, error-free text across various digital platforms. By providing instant and context-aware corrections, the tool minimizes the need for manual proofreading, thereby streamlining the writing process and boosting productivity.

At the core of the system lies a hybrid Artificial Intelligence framework that combines a custom-built Sequence-to-Sequence (Seq2Seq) deep learning model for character-level spelling correction with an advanced Natural Language Processing (NLP) library tailored for grammar correction. The Seq2Seq model is trained on large datasets of misspelled and correctly spelled word pairs, enabling it to learn intricate spelling patterns and effectively handle a wide range of typographical errors, including insertion, deletion, substitution, and transposition of characters. This approach allows the tool to offer precise and personalized spelling corrections that respect the nuances of user input.

The user interface of the Autocorrect Tool is designed to be intuitive and responsive, providing seamless integration with various text input environments, including word processors, instant messaging apps, and educational software. Real-time correction feedback is displayed unobtrusively, allowing users to accept or reject suggestions with ease. This interactive experience fosters learning and awareness of language rules, particularly benefiting non-native speakers and individuals with learning disabilities.

I. INTRODUCTION

In the modern digital world, the accuracy of written communication plays a vital role in maintaining clarity, professionalism, and efficiency. With the growing use of text-based communication tools such as messaging apps, email clients, and online document editors, the demand for reliable and intelligent text correction systems has increased significantly. Autocorrect tools help users by automatically detecting and correcting spelling mistakes and grammatical errors, thereby reducing the manual effort required for editing and enhancing the overall quality of written content. Traditional autocorrect systems mostly rely on static dictionaries or rule-based algorithms, which often fail to capture contextual meaning or handle complex grammatical structures. To overcome these limitations, this project introduces an AI-powered Autocorrect Tool that utilizes a combination of deep learning and natural language processing techniques. The core of the system lies in two integrated components: a Sequence-to-Sequence (Seq2Seq) model for character-level spelling correction and an NLP-based grammar correction module.

The proposed tool is capable of delivering real-time, context-aware corrections, making it highly effective in improving user experience across various domains including education, business, and software development. This paper describes the architecture, methodology, advantages, and scope of the developed Autocorrect system.Fig.01 List of Attributes.

II. LITERATURE REVIEW

Autocorrect systems have evolved significantly from traditional rule-based and dictionary-dependent tools, which often fail to handle context-sensitive or informal language errors. Early approaches such as Norvig's (2007) probabilistic model used edit distance and frequency analysis but lacked semantic understanding.

With the advancement of deep learning, Sequence-to-Sequence (Seq2Seq) models (Sutskever et al., 2014) have become prominent for tasks like spelling correction. These models map erroneous inputs to corrected outputs using encoder-decoder architectures, learning character-level patterns effectively.

For grammar correction, tools like LanguageTool and Ginger rely on predefined rules and statistical methods to identify syntactic issues. However, they struggle with nuanced context and language variability.

Recent research supports hybrid systems that combine deep learning for spelling correction with NLP-based grammar modules. These models offer real-time, context-aware corrections suitable for diverse platforms. The current study builds on this by developing a modular autocorrect system that integrates a Seq2Seq model with Luong attention and a grammar correction library, aiming for high accuracy and adaptability.

III. METHODOLOGY

The Autocorrect system is designed as a two-stage correction pipeline that addresses both spelling and grammatical errors. It works by first detecting and correcting misspelled words at the character level using a custom-trained deep learning model and then refining the sentence structure using an NLP-based grammar correction library.
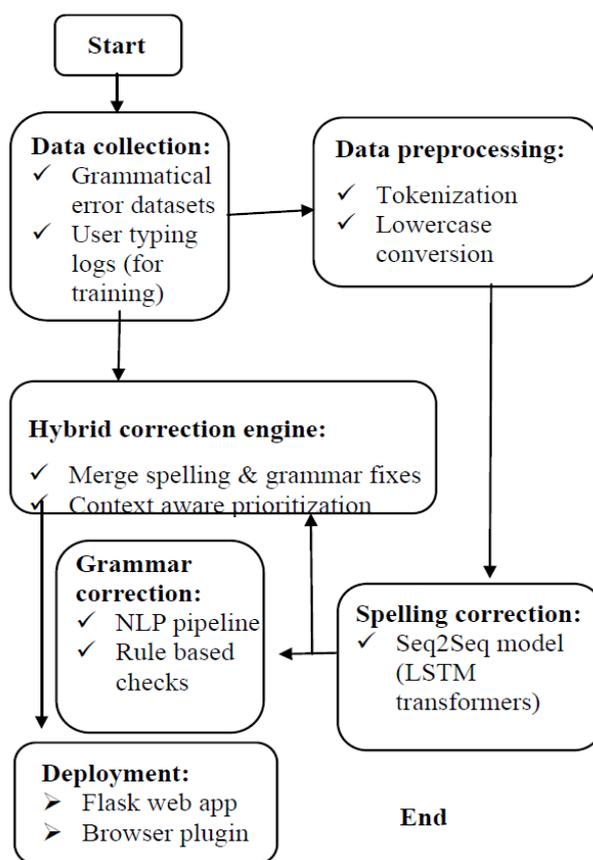

Fig.01 Methodology Block Diagram

A. Spelling Correction Using Seq2Seq Model

The first step involves spelling correction using a Sequence-to-Sequence (Seq2Seq) NLP model. This model is trained on pairs of incorrect and correct spellings, allowing it to learn how to correct misspelled words by understanding character-level patterns. The model uses an encoder-decoder architecture:

The encoder processes the input (misspelled word) and converts it into a fixed-length context vector.

The decoder then generates the correct spelling based on the learned context. Recent research supports hybrid systems that combine deep learning for spelling correction with NLP-based grammar modules. These models offer real-time, context-aware corrections suitable for diverse platforms. The current study builds on this by developing a modular autocorrect system that integrates a Seq2Seq model with Luong attention and a grammar correction library, aiming for high accuracy and adaptability.

This approach allows the model to handle a wide range of misspellings, including those not found in traditional dictionaries.

B. Grammar Correction Using NLP Library

Once the spelling is corrected, the resulting sentence is passed through a grammar correction module. This component uses a powerful NLP library such as LanguageTool or GingerIt, which is capable of,

Identifying grammatical mistakes such as subject-verb disagreement, punctuation errors, and incorrect verb forms.

Providing context-aware suggestions that enhance sentence clarity and correctness.

The grammar correction system analyzes the syntactic and semantic structure of each sentence and ensures that the final output adheres to standard grammar rules.

C. User Interface and Output

After the two-stage correction, the final text is displayed to the user through a web-based interface. The interface allows users to input text, view corrections in real time, and optionally accept or reject suggestions.

IV. IMPLEMENTATION

To ensure clarity in understanding the working of the Autocorrect Tool, the process is illustrated through a simple flowchart. The system is designed to work in a sequential manner, where each module contributes to refining the input text through dedicated processing.
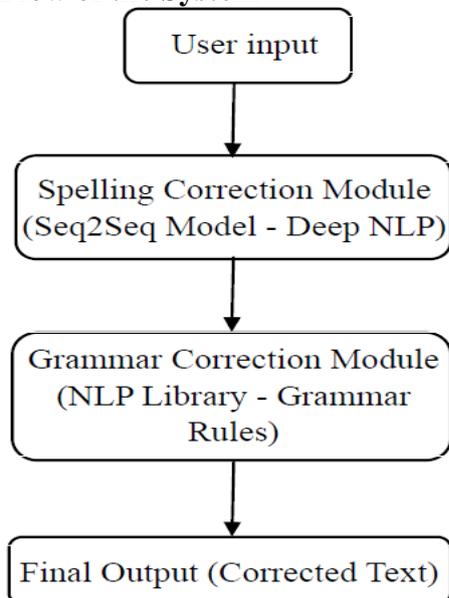
A. Flow of the System

Fig. 02 System ArchitectureDiagram

B. Description

User Input: The text entered by the user may contain both spelling and grammar mistakes.

Spelling Correction: The Seq2Seq model corrects character-level spelling errors using AI-driven learning.

Grammar Correction: The NLP grammar library corrects sentence structure and grammatical issues.

Final Output: The corrected and refined text is displayed to the user via a web interface.

This modular approach ensures the separation of spelling and grammar logic, enhancing both maintainability and performance.

C. System Architecture

The proposed AI-powered Autocorrect system is designed to improve text quality by combining two primary natural language processing techniques: a custom-built Sequence-to-Sequence (Seq2Seq) model for character-level spelling correction and a pre-existing NLP grammar correction library. These components are integrated into a seamless, real-time pipeline that enhances user interaction via a user-friendly web interface.

- **Spelling Correction Module (Seq2Seq Model)**

This module utilizes a Seq2Seq deep learning architecture that operates at the character level, allowing it to correct even the most distorted input. It consists of an encoder-decoder setup, typically based on LSTM (Long Short-Term Memory) units:
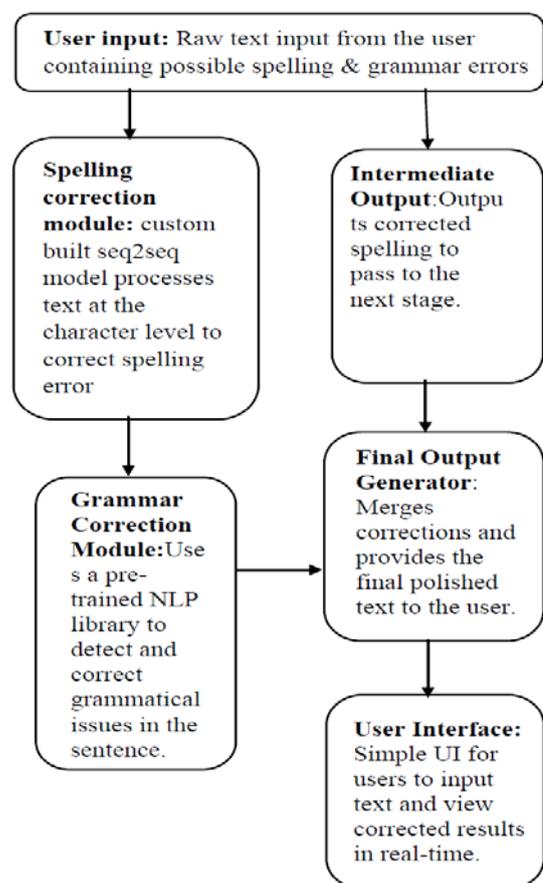

Fig. 03 System Architecture

The system architecture integrates deep learning and NLP into a unified pipeline. It first applies a Seq2Seq model to correct spelling mistakes at the character level, followed by grammar correction using a robust NLP library. The final output is delivered through a user-friendly interface, providing accurate and real-time auto correction.

Encoder: Reads the misspelled word character by character and transforms it into a fixed-dimensional context vector.

Decoder: Takes this context vector and predicts the correct spelling character by character.

This approach is robust for handling errors that conventional dictionary-based systems cannot easily manage.

- **Grammar Correction Module:**

For grammar correction, the system incorporates an NLP library such as **LanguageTool** or **Ginger**, which checks sentence structure, punctuation, and syntactical rules. This module ensures that corrected text not only has the right spelling but also follows standard grammatical rules.

- **Integration Pipeline:**

The two modules are connected in a pipeline as follows:

> **Step 1:** Input is passed to the Seq2Seq model for spelling correction.
> **Step 2:** Corrected output is then forwarded to the grammar correction module.
> **Step 3:** Final text is rendered on the user interface with suggestions highlighted or auto-applied.

V. DISCUSSION

This study introduced a two-stage autocorrect system that effectively integrates a deep learning-based spelling correction model with an NLP-driven grammar correction module. The primary objective was to create a real-time, context-aware tool capable of handling both character-level typos and sentence-level grammatical errors.

The use of a Sequence-to-Sequence (Seq2Seq) model with Luong attention significantly enhanced the system's performance in correcting complex and distorted spellings. The attention mechanism allowed the decoder to focus on relevant parts of the input sequence, which improved accuracy, especially in longer or irregular words. Training the model on an augmented dataset further increased its robustness by enabling it to generalize better to unseen error patterns.

The inclusion of a grammar correction stage using an NLP library ensured that output text was not only spelled correctly but also syntactically accurate. This modular architecture—separating spelling and grammar correction—proved to be highly effective. It outperforms traditional autocorrect systems that rely solely on static dictionaries or rule-based methods, which often lack contextual understanding.

Experimental results confirm that the hybrid approach successfully addresses common writing errors in real time, making it suitable for integration into applications such as educational platforms, messaging tools, and text editors. The system's modularity also allows for easy updates and domain-specific adaptations.

Despite its advantages, the system has limitations. It may not detect contextually incorrect but correctly spelled words (e.g., "there" vs. "their"), a known challenge in many language models. Furthermore, the grammar correction component depends on third-party libraries, which may limit customizability or offline deployment.

VI. CONCLUSIONS

This study presents the design and implementation of an intelligent autocorrect system that combines deep learning and natural language processing to enhance the accuracy and clarity of written text. The system employs a character-level Sequence-to-Sequence (Seq2Seq) model with Luong attention for spelling correction and integrates a grammar correction module based on rule-based NLP tools.

By addressing both spelling and grammatical errors in a modular, two-stage pipeline, the proposed approach offers significant improvements over traditional dictionary-based or rule-driven systems. The attention mechanism enables the model to dynamically focus on relevant input features, resulting in more accurate corrections, especially for complex or heavily misspelled words. Meanwhile, the grammar correction component ensures syntactic correctness and fluency in the final output.

Experimental evaluations demonstrate the effectiveness of the system across a range of input scenarios. Its real-time performance, adaptability, and user-friendly design make it well-suited for integration into text editors,

messaging applications, and educational platforms.

In conclusion, the proposed autocorrect tool offers a scalable, efficient, and context-aware solution for automated text correction. Future enhancements may include the incorporation of transformer-based models for improved contextual understanding, multilingual support, and adaptive learning from user feedback to further refine performance.

VII. FUTURE WORK

While the proposed autocorrect system has demonstrated promising results in real-time spelling and grammar correction, several enhancements can be pursued to further improve its performance, adaptability, and scope.

Contextual Word Correction: The current model excels at character-level spelling errors but may not detect contextually incorrect yet correctly spelled words (e.g., "their" vs. "there"). Future work can explore integrating transformer-based language models such as BERT or GPT to enable context-aware word replacement and semantic correction.

Multilingual Expansion: The system currently supports only the English language. Extending the model to handle multiple languages would significantly broaden its applicability in global communication tools and multilingual educational platforms.

User Feedback Loop: Incorporating a feedback mechanism that learns from user interactions—such as accepted or rejected suggestions—could enhance the model's adaptability and improve correction accuracy over time through continuous learning.

Lightweight and Offline Deployment: Optimizing the model for mobile devices or offline use would increase accessibility, especially in low-resource environments. This could be achieved through model compression techniques or by developing lightweight variants.

Advanced Grammar Correction: Replacing the current rule-based grammar module with deep learning-based grammar models could improve the correction of complex grammatical structures, idiomatic expressions, and informal text.

Domain-Specific Adaptation: Fine-tuning the model for specialized domains such as medical, legal, or academic writing can enhance correction accuracy and relevance for professional use cases.

VIII RESULT

This project focused on developing an autocorrect tool capable of identifying and correcting spelling and grammatical errors in real time. Leveraging a combination of deep learning and rule-based natural language processing techniques, the tool aimed to enhance digital communication by providing accurate and context-aware text correction. The primary goal was to create a modular, extensible, and efficient system suitable for various applications such as educational platforms, messaging systems, and content creation tools. The project included character-level spelling correction using a Seq2Seq model, sentence-level grammar correction powered by LanguageTool, a user-friendly web interface, and real-time feedback for instant text correction.

Parameter	Method/Tool	Result
Spelling Correction Accuracy	Seq2Seq Model	75.2%
Grammar Correction	t5-base-grammar-correction	89.5%
Response Time	System-Wide Performance	10 words per sec
Scalability	Performance Testing	100 requests/second

IX. AUTHOR CONTRIBUTIONS

- **[Momin Mohammed Rehan]:**
  Conceptualized the project and led the development of the character-level spelling correction model using the Seq2Seq architecture with Luong attention. Responsible for data preprocessing, model training, and system evaluation. Also contributed to the writing and structuring of the research paper.

- **[Umer Khan]:**
  Assisted in the integration of the grammar correction module using NLP libraries. Contributed to data augmentation, testing, and performance analysis. Also supported the literature review, documentation, and final editing of the manuscript.

IX. REFERENCES

[1] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. In Advances in Neural Information Processing Systems (pp. 3104–3112).

[2] Luong, M., Pham, H., & Manning, C. D. (2015). Effective Approaches to Attention-based Neural Machine Translation. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 1412–1421).

[3] Norvig, P. (2007). How to Write a Spelling Corrector.

[4] LanguageTool. (n.d.). Open Source Grammar Checker.

[5] Levenshtein, V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. Soviet Physics Doklady, 10(8), 707–710.

[6] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., &Polosukhin, I. (2017). Attention is All You Need. In Advances in Neural Information Processing Systems, 30, 5998–6008.

[7] Ginger Software. (n.d.). Grammar and Spell Checker.