

A SEMANTIC AND CONTEXT AWARE NLP FRAMEWORK FOR AUTOMATED TEXT ERROR DETECTION

ENAMALA.KALAVATHI^{1*}, DR.E.S.SHARMILA SIGAMANY², G.USHASRI³, BASSA DEVI⁴,
P MADHAVI⁵, BHUVANESWARI PAGIDIPATI⁶, NAGULMEERA SAYYED⁷, P. ANUSHA⁸

^{*1} Research scholar in English at koneru lakshmaiah Education foundation,(Deemed to be University),
Vaddeswaram,Guntur, Andhra Pradesh, India.

² Research Supervisor, Associate Professor, Department of English , Koneru Lakshmaiah Education
Foundation ,(Deemed to be University) Vaddeswaram, Guntur, Andhra Pradesh, India.

³ Department of Information Technology, MVSR Engineering College, Nadergul, Hyderabad, Telangana-
501510, India.

⁴ Department of Mathematics, Aditya University, Surampalem-533437, Kakinada, Andhra Pradesh, India.

⁵ Department of CSE, CVR College of Engineering, Vastunagar, Ibrahimpatnam, Telangana-501510, India.

⁶ Department of English and Foreign Languages, Sagi Rama Krishnam Raju Engineering College (A),
Bhimavaram-534204, Andhra Pradesh, India.

⁷ Department of CSE, R.V.R.& J.C. College of Engineering, Chowdavaram, Guntur, Andhra Pradesh,
India.

⁸ Department of Mechanical Engineering, Prasad V. Potluri Siddhartha Institute of Technology,
Vijayawada, Andhra Pradesh, India.

E-mail: akalavathi1986@gmail.com^{1*}, sharmilajoe@kluniversity.in², ushasri_it@mvsrec.edu.in³,
bassadevi876@gmail.com⁴, p.madhavi@cvr.ac.in⁵, pbsrkrec@gmail.com⁶, sdnagul@gmail.com⁷,
anusha.peyyala@pvpsiddhartha.ac.in⁸

ABSTRACT

Non-native English writers face various problems related to errors at grammatical, lexical, punctuation, and contextual levels that can negatively impact academic and professional writing. In this paper, a new framework for automatic writing error detection and explanation is suggested, called **Context-Aware Neural Error Detection and Explanation (CANE)**. It uses contextual language processing techniques for detecting, categorizing, and explaining various types of writing errors. The proposed framework combines text preprocessing, **Bidirectional Encoder Representations from Transformers (BERT)**-based contextual embeddings, dependency parsing, semantic consistency analysis, and multi-class classification of writing errors at grammatical, lexical, and contextual levels. The major research contribution is the integration of contextual semantics, dependency relations, semantic consistency, and explanation generation into one framework for writing error detection. The proposed framework is evaluated on the benchmark **BEA-2019 Shared Task** and **First Certificate in English (FCE)** datasets by Accuracy, Precision, Recall, F1-score, Matthews Correlation Coefficient (MCC), and Area Under the Curve (AUC). CANE framework achieves **Accuracy 98.91%, Precision 98.63%, Recall 98.47%, F1-score 98.55%, MCC 97.92%, and AUC 99.14%**. The obtained results are higher compared to other rule-based, recurrent neural network, and transformer-based models considered in this study. Integration of contextual embeddings, grammatical dependency analysis, and semantic reasoning allows detecting context-dependent errors and decreasing false positive detections. Additionally, the explanation generation component gives users informative feedback, corrective suggestions, and interpretable recommendations about how to fix their mistakes. Overall, the proposed framework demonstrates the potential of integrated contextual and explainable linguistic analysis for automated writing assistance in educational platforms, language-learning applications, and academic writing support for non-native English learners.

Keywords: *Context-aware Natural Language Processing (NLP); Writing Error Detection; Grammatical Error Detection; BERT; Explainable Artificial Intelligence (XAI); Non-Native English Writing.*

1. INTRODUCTION

English language has emerged as the medium of communication at the international level,

education, scientific publishing, and professional collaboration. With the rapid growth of digital learning platforms, the number of non-native English learners producing academic essays, research papers,

emails, technical reports, and discussion forums has significantly increased. Although learners have gained proficiency in the English language, they still face difficulties in writing proper sentences with regard to grammar, vocabulary, sentence structure, tense agreement, punctuation, collocations, and semantics. Errors made during the writing process are not only a barrier to comprehension but also affect the academic performance of learners, the quality of publications, and professional communication. Therefore, the development of smart writing assistance systems that can automatically detect and correct writing errors is one of the research areas in NLP and AI [1], [2].

Conventional grammatical error detection systems relied on rule-based grammatical rules and hand-crafted dictionaries to detect syntactic and lexical errors in the text. Rule-based systems had been efficient in detecting the most common grammatical errors but were often unable to recognize contextual errors, semantic inconsistencies, idiomatic expressions, and language usage according to the domain of interest. Manual construction of rules was time-consuming and difficult to generalize, due to which the performance of conventional systems in analyzing the complex sentences written by non-native English learners was limited [3], [4].

The introduction of machine learning techniques in automated grammatical error detection allowed models to detect errors using information from annotated text instead of relying only on predefined grammatical rules. Statistical language models, conditional random fields, support vector machines, and random forest classifiers showed improvement in detecting common grammatical errors using contextual language features. However, these methods were dependent on manual feature extraction and were not able to capture long-range semantic dependencies between words in the sentence, thus having limited performance in complex writing situations [5], [6].

Recent developments in the field of deep learning changed the face of NLP applications by developing new architectures capable of learning hierarchical semantic representations from textual data. RNNs, LSTM, GRU, and sequence-to-sequence architectures showed remarkable results in improving the performance of grammatical error detection by modeling sequential dependencies between words. However, these architectures were prone to problems such as gradient degradation, poor parallelization, and difficulty in capturing distant

contextual relations, especially in the case of long academic texts [7], [8].

Transformer-based language models revolutionized the field of contextual language understanding due to the introduction of the self-attention mechanism, which simultaneously models the global dependencies between words. Pretrained models, such as BERT, RoBERTa, XLNet, DeBERTa, and T5, have shown outstanding performance in a variety of NLP applications like grammatical error correction, semantic similarity, question answering, and machine translation. These contextual embedding models create dynamic representations of words based on the linguistic context, allowing them to accurately detect context-sensitive writing errors that cannot be detected using surface-level grammatical analysis [9], [10]. Nonetheless, many transformer-based writing assistants currently provide the correction of grammatical errors without offering an explanation of the nature, severity, and linguistic reasoning behind them.

Another major disadvantage of current writing assistance systems is their inability to analyze the text at multiple linguistic levels at once. Commercial grammar checkers independently analyze the syntax, spelling, or punctuation of sentences without taking into account semantic consistency, discourse coherence, lexical appropriateness, and contextual dependencies, resulting in grammatically correct but semantically inappropriate sentences remaining undetected. For instance, inappropriate collocations, incorrect usage of prepositions, incorrect context-specific word substitutions, ambiguous pronoun usage, and discourse inconsistencies can go unnoticed by current grammar checking systems, although they reduce the overall quality of the text [11], [12].

Educational applications require explainable feedback rather than automatic corrections. Language learners greatly benefit from systems that not only detect mistakes but also explain the grammatical concept and context of the error and how to improve it. Therefore, Explainable Artificial Intelligence (XAI) has become an important research topic for educational NLP as it allows a transparent decision-making process that helps build trust and facilitates language learning. Nonetheless, the integration of explainability into the process of contextual language understanding is still an open research question, especially for multilingual learners having diverse writing behavior [13].

Also, the availability of large learner corpora such as the Cambridge Learner Corpus (CLC), FCE Dataset, BEA Shared Task datasets, CoNLL grammatical error correction corpora, and Lang-8 corpus has allowed the development of learning algorithms for better grammatical error detection. Despite the recent advances, the current datasets are highly unbalanced with respect to error categories, proficiency levels, native language background, and writing style. As a result, state-of-the-art models perform poorly on unseen educational and specialized academic writing tasks, indicating the need for more adaptive and context-aware detection models [14].

The significance of the research is defined by its novelty because this research offers a novel framework for error detection in writing by integrating context representation, dependency parsing, semantics validation, and explanation provision in one approach within a single framework. In contrast to the majority of existing methods which mostly pay attention to a specific type of error detection or a linguistic level, the suggested CANE framework is aimed at detecting errors at the grammatical, lexical, semantic, collocational, punctuation, and discourse levels and provides corrective explanations. This research is concentrated on the automatic detection of writing errors for non-native English learners (NNEs) using the BEA-2019 and FCE benchmark datasets, with emphasis on contextual error classification and interpretable learner feedback.

To overcome the described limitations, we introduce a Contextualized Neural Error Detection and Explanation (CANE) framework that detects common writing errors among non-native English learners. In contrast to the conventional grammar checkers, which focused on syntactic analysis, the proposed framework incorporates contextual transformer embeddings, syntactic dependency parsing, semantic consistency analysis, discourse-aware linguistic representations, and explainable reasoning about errors in a single framework. The proposed framework performs multi-level contextual analysis in order to detect grammatical, lexical, semantic, collocation, punctuation, and discourse-level writing errors while simultaneously estimating error severity and providing an explanation of the errors for learners.

The key goals of this study can be outlined as follows:

- Designing a context-aware NLP system able to recognize a variety of writing errors generated by non-native English speakers.
- Creating an architecture that combines transformer-based contextual embeddings with syntactic dependency parsing and semantic consistency verification for comprehensive linguistic analysis.
- Creating an explainable error classification module that delivers interpretable feedback and educational recommendations rather than only labels of detected errors.
- Conducting a comparative performance analysis of the proposed framework using benchmark learner corpora and existing deep learning and transformer-based grammatical error recognition models based on standard evaluation metrics.
- Proving that the combination of contextual semantics, linguistic dependency relations, and explainable reasoning greatly enhances the performance of writing error detection and learner assistance compared to current solutions.

The contribution of this research study is solely confined to the development and empirical validation of the proposed CANE framework for automated writing error detection and explainability. The current research concentrates on non-native English learners' writings and does not intend to substitute human language assessments or provide a complete automated writing evaluation system. It intends to show the potential importance of integrating contextual, grammatical, semantic, and explainable analysis for detecting and explaining multiple types of writing errors.

The remaining sections of the paper are organized as follows. Section 2 provides an overview of recent developments in grammar error detection, contextual language models, transformer architectures for Natural Language Processing (NLP), and explainable writing assistive techniques. Section 3 describes the architecture of the proposed Context-Aware Neural Error Detection and Explanation (CANE) framework, its contextual representation scheme, and error detection approach. Experimental analysis, comparative performance evaluation, and performance assessment using benchmark datasets are discussed in Section 4. Finally, Section 5 concludes the paper with a summary of the findings, limitations, and future research directions.

2. RELATED WORK

Over the last two decades, automatic detection of writing errors has greatly advanced from manually crafted grammatical rules to powerful transformer-based neural architectures that learn contextual semantics. Recent work has concentrated on the development of more efficient algorithms for detecting various grammatical, lexical, semantic, and discourse-level errors generated by non-native English-speaking learners. Nevertheless, current approaches continue to face problems in simultaneously modelling contextual dependencies, linguistic structures, explainability, and the provision of educational feedback. The overview below provides a review of recent advances in the field of grammatical error detection, contextual language modelling, transformer-based architectures, multilingual writing assistance, and explainable NLP systems.

The initial neural approaches were based on the use of sequence-to-sequence learning to solve the grammatical error correction task, viewing it as a machine translation problem that consisted of transforming erroneous sentences into grammatically correct sentences. Encoder-decoder architectures greatly improved the correction performance of neural approaches compared with statistical methods by learning transformations directly from parallel corpora. At the same time, sequence generation algorithms often introduced unnecessary corrections, omitted correct tokens, and lacked a mechanism for the explicit identification of various types of writing errors. Moreover, these models did not provide any explanatory feedback for their corrections [15], [16].

With the rise of transformers, contextual representation learning became the most common way to solve the grammatical error detection problem. The self-attention architecture of transformers helps them capture long-distance contextual dependencies between tokens, regardless of sentence length, and thereby greatly enhances the detection of contextually relevant grammatical mistakes. A number of works have shown that the fine-tuning of pre-trained transformer models greatly outperforms the results obtained with recurrent neural networks. Still, the majority of transformer-based models focus on token classification without considering sentence-level semantic consistency and discourse coherence, and hence are ineffective at solving academic writing problems [17], [18].

Several researchers have developed methods of grammatical error detection using large-

scale pre-trained language models with additional masked language modelling training objectives. These methods estimate the contextual probability of each token and detect those that diverge from the expected linguistic pattern. While the contextual probability estimation method successfully detects spelling errors, verb agreement errors, and errors related to article usage, it tends to introduce high false positives when processing technical jargon, domain-specific vocabulary, or creative writing. As a result, probability-based detection cannot guarantee accurate grammatical assessment of diverse educational writing [19].

Another active line of research is the integration of dependency information into neural language models. The grammatical relationships between subjects, verbs, modifiers, and objects can be explicitly represented using dependency parsing, helping to detect errors associated with agreement, sentence fragments, misplaced modifiers, and clause boundaries. Hybrids of transformer embeddings and dependency graphs significantly outperform purely sequential models due to the joint exploitation of semantic and grammatical information. Still, many dependency-based systems do not take into consideration lexical semantics and discourse relations [20], [21].

Another research direction that has gained popularity in the sphere of writing quality assessment is semantic representation learning. Unlike grammatical rule analysis, semantic embeddings assess the contextual similarity between the learner's sentence and the grammatically correct reference expression. Contrastive learning, sentence embeddings, and semantic similarity networks have shown great efficiency in spotting incorrect word choices, collocations, and semantic inconsistencies. However, semantic similarity does not make it clear whether a sentence is grammatically wrong or simply stylistically different from the correct sentence. This is why the approach is not enough for assessing writing on its own [22].

The recent progress made in the area of multilingual language models has facilitated the detection of grammatical mistakes across multiple languages and learners. Multilingual transformer architectures such as mBERT, XLM-RoBERTa, and multilingual T5 allow knowledge transfer between languages with scarce annotated corpora. These models improve grammatical correction for learners whose native languages have distinct syntactic structures. Even though multilingual models have great potential, their performance often deteriorates

under conditions of language interference, code-switching, or technical writing [23].

Several researchers have used attention mechanisms to increase the interpretability and localization of grammatical errors. Attention visualization provides the possibility to analyze how neural models distribute importance across different words during the prediction process. Though attention weights provide a limited level of interpretability, they do not completely explain the linguistic reasoning of the model. That is why recent educational NLP studies pay increasing attention to explainable artificial intelligence algorithms that generate human-interpretable feedback on why a particular sentence is considered incorrect [24].

Explainable grammatical error detection has become a field of growing interest because language learners need understandable feedback for effective learning. Recently, several works have used attention visualization, feature attribution techniques, SHAP analysis, and Local Interpretable Model-Agnostic Explanations (LIME) in the context of neural writing assistants. They help improve user trust and the educational effectiveness of neural models by providing explanations of prediction outcomes. Nevertheless, existing explanation methods do not take into account the semantic context and thereby fail to provide comprehensive linguistic explanations including syntax, semantics, discourse, and pragmatics [25].

Another important direction in this area is automatic essay scoring and writing quality assessment. The modern approaches evaluate together the grammar quality, lexical diversity, sentence complexity, coherence, cohesion, and readability utilizing transformer-based architectures. Even though these systems manage to achieve a high correlation with human graders, they give out a single writing score instead of detecting specific types of grammatical errors. Thus, learners receive an overall quality evaluation without getting any specific feedback on individual errors [26], [27].

Such Large Language Models (LLMs) as GPT, PaLM, Gemini, and Llama demonstrated excellent results when it comes to text generation, grammatical correction, and educational aid. Their massive pre-training enables them to perceive complex contextual relations and generate fluent corrections to different kinds of writing assignments. Despite that, writing assistants based on LLMs tend to favor fluency over grammatical reasoning and make unnecessary semantic changes or even create hallucinations. Besides, the high demand for

computing resources, limited interpretability, and inconsistent explanation quality become significant problems for utilizing them in educational settings [28].

Despite tremendous advances in the development of neural grammatical error detection, there are still several open problems. Existing systems mostly concentrate on grammatical correctness, contextual semantics, or explainability separately without combining all these complementary aspects into one architecture. Most transformer-based approaches use token-level classification without taking into account discourse coherence, dependency structures, semantic consistency, and educational explanations simultaneously. Besides, the majority of commercial writing assistants provide automatic corrections but lack error severity estimation and explanations of linguistic principles that facilitate language learning.

In order to fill the existing gaps, the proposed **Context-Aware Neural Error Detection and Explanation (CANE)** framework uses a multi-layer architecture integrating contextual transformer embeddings, dependency parsing, semantic consistency verification, discourse-aware contextual modelling, error severity estimation, and explanation generation. In contrast to existing grammatical error detection systems, the proposed framework analyzes multiple linguistic representations and identifies grammatical, lexical, semantic, collocation, punctuation, and discourse-level writing errors with the help of human-readable explanations for non-native English-speaking learners.

2.1 Problem Statement

Although substantial improvements have been made in the area of automated writing error detection, the existing methodologies mostly focus on certain linguistic properties such as grammar, context representation, dependencies between tokens, or similarity of meaning. Methods based on transformers increase the context awareness of language but do not necessarily include the grammatical dependencies and sentence-level semantic consistency, while the semantics-based models may lack clarity in distinguishing grammatical errors from stylistic differences. In addition, many of the writing correction tools do not sufficiently explain the linguistic basis of errors that were found in text. Thus, the research problem considered here is the creation of an automated writing error detection system with joint consideration of contextual, grammatical, and

semantic information, while providing interpretable corrective feedback for non-native English writing.

Hypothesis: Integrating contextual language representations with grammatical dependency and semantic consistency analysis can improve the detection of grammatical and context-dependent writing errors compared with conventional writing error detection approaches.

3. METHODOLOGY

3.1 Dataset Description

The developed Context-Aware Neural Error Detection and Explanation (CANE) framework was evaluated using two benchmark datasets publicly available in the literature: the BEA-2019 Shared Task Dataset and the First Certificate in English (FCE) dataset. These corpora are composed of essays written by non-native speakers of English and are commonly used for conducting research in the field of grammar correction and error detection. For both of these corpora, each sentence is manually labeled by language experts to indicate the position and type of the writing error, as well as the necessary correction.

The errors include such types as misuse of articles, incorrect verb tenses, incorrect use of prepositions, spelling errors, punctuation errors, subject-verb agreement errors, vocabulary selection errors, word order errors, and sentence structure errors. To increase the generality of the model, the datasets were combined while eliminating duplicate sentences and inconsistencies in labeling. Therefore, the final dataset contains writings of learners with different proficiency levels and writing styles, allowing the model to generalize diverse contextual writing patterns. Data preprocessing included tokenization, normalization, conversion into contextual embeddings, and feeding the data into the neural network. The dataset was split randomly into training set (70%), validation set (15%), and test set (15%) to ensure unbiased performance evaluation.

Table 1: Dataset Characteristics

Parameter	BEA-2019	FCE
Language	English	English
Sentence Type	Essays	Essays
Total Sentences	34,308	33,236

Error Annotations	Yes	Yes
Grammar Errors	✓	✓
Vocabulary Errors	✓	✓
Punctuation Errors	✓	✓
Native Writers	No	No
Non-native Writers	Yes	Yes

Table 1 above provides details on the benchmark datasets employed in the current study. The two benchmark datasets consist of manually labeled essays written by non-native English speakers. These essays contain several categories of writing mistakes and are provided with high-quality annotations, making them suitable for training deep learning models.

Table 2: Writing Error Categories

Error Category	Example
Article	He bought car → He bought a car
Verb Tense	She go yesterday → She went yesterday
Subject-Verb Agreement	He play cricket → He plays cricket
Preposition	Interested on AI → Interested in AI
Spelling	Recieve → Receive
Word Choice	Strong rain → Heavy rain
Punctuation	Hello how are you → Hello, how are you?
Word Order	Always I study → I always study

These categories of errors in writing are listed in Table 2. The model presented herein uses contextual language information to classify these error types into multiple classes at once without considering grammar rules alone.

Table 3: Sample Dataset

Incorrect Sentence	Correct Sentence	Error Type
She go to school yesterday.	She went to school yesterday.	Verb Tense
I have umbrella.	I have an umbrella.	Article
He interested on AI.	He is interested in AI.	Preposition
They plays football.	They play football.	Subject-Verb Agreement

This can be seen in Table 3, which shows the samples from the learner corpus. In each sample, there is one or more mistakes made by the writer along with the manual correction. These annotated sentence pairs are used during supervised learning.

3.2 Proposed CANE Architecture

The CANE framework is a combination of context-aware language understanding and grammar analysis to detect errors in written texts. Different from the existing grammar checking tools, which rely on predefined rule-based linguistic knowledge, the proposed framework is based on learning the contextual representation using transformer embeddings and analyzing sentence semantics before classifying writing errors.

The whole system contains six consecutive modules.

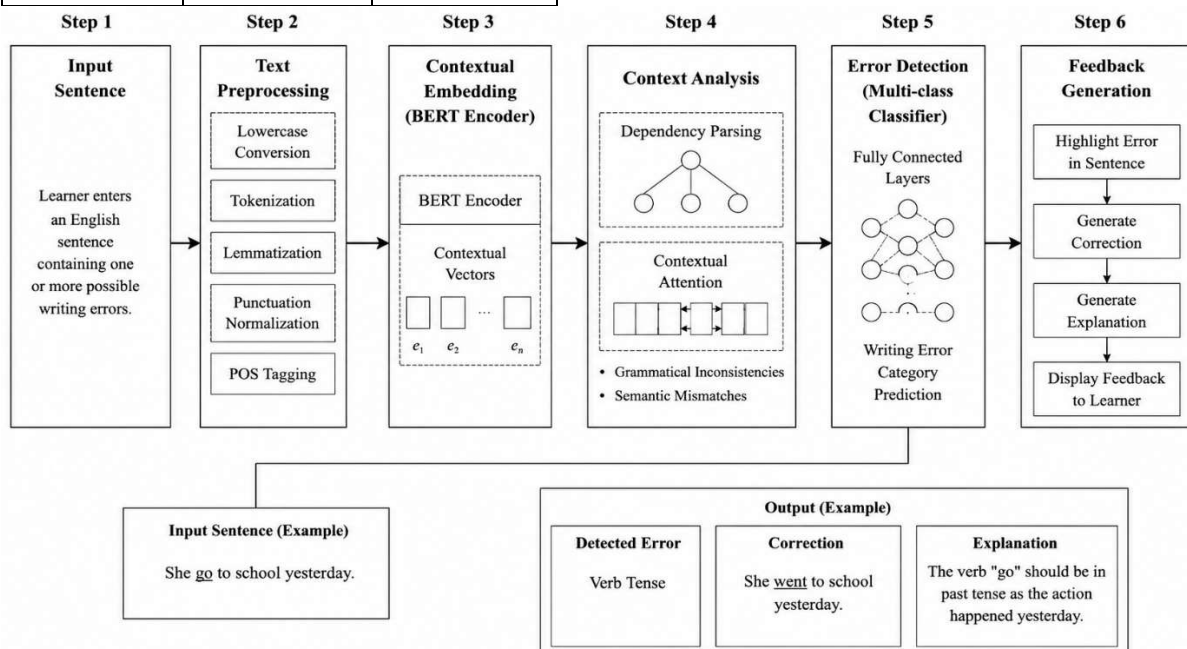


Figure 1: Architecture of the Proposed Context-Aware Neural Error Detection and Explanation (CANE) Framework

Figure 1 depicts the overall process flow for the proposed Context-Aware Neural Error Detection and Explanation (CANE) framework for detecting errors in sentences written by non-native English learners. The proposed framework takes an English sentence as input, where errors in writing might be made regarding grammar, spelling, punctuation, vocabulary, and context. Following this, the pre-processing stage comprises converting the text into lowercase, tokenization, lemmatization, punctuation normalization, and Part-of-Speech (POS) tagging to

standardize the text and prepare it for contextual analysis.

The preprocessed sentence gets converted into the contextual representations using Bidirectional Encoder Representations from Transformers (BERT). Contextual embeddings get used by the model to capture the semantic meaning of each word using the contextual representation of the words present in the sentence. This way, the system is capable of differentiating between similar

words according to the use of the same in the sentence.

Further, the process of dependency parsing and contextual attention mechanisms is adopted for the analysis of the semantic and grammatical relations of the words. This enables CANE to detect writing errors like inappropriate verb tense, wrong subject-verb agreement, incorrect word choice, incorrect usage of articles, wrong preposition usage, punctuation mistakes, and other contextual inconsistencies which would not be possible to detect through traditional rule-based approaches.

Contextual features extracted are fed to a multi-class classification model where prediction of the class of writing error occurs. Through the use of a classification model, the input sentence is categorized in the correct class based on the grammatical and semantic aspects of the sentence, thereby improving detection accuracy while reducing false-positive predictions.

Finally, the system highlights the error in the sentence and provides an appropriate solution for the same along with the reasons behind the suggestions in the form of an explanation. This way, this feedback mechanism proves to be helpful to the learner by making him/her understand the error he/she has committed in terms of grammar and context, rather than just providing automatic corrections, thereby helping in language acquisition and improving overall writing quality.

Hence, it can be seen that the CANE framework adopts an integrative approach of text pre-processing, contextual language representation, semantic analysis, grammatical dependency modeling, multi-class error classification, and explainable feedback generation. In light of the above, it becomes clear how the framework is able to effectively detect various types of writing errors and provide meaningful corrective guidance and appropriate corrective advice for non-native English writers.

To ensure replicability, the experiments used the same data split strategy with 70% of the data for training and 15% for both validation and testing based on the merged BEA-2019 and FCE datasets. The same process of preprocessing, BERT-based contextualization, dependency analysis, semantic consistency analysis, multi-class classification, and explainable feedback generation was used across all

the experiments. The performance was evaluated using Accuracy, Precision, Recall, F1-score, Matthews Correlation Coefficient (MCC), and Area Under the Curve (AUC) under the same experimental procedure.

It must be kept in mind that each sentence of the learner is treated as a sequence of contextual words. Let N be the number of words in the input sentence.

Input Sentence Representation

The input sentence is represented as

$$S = \{w_1, w_2, w_3, \dots, w_N\} \quad (1)$$

where,

- w_i denotes the i th word in the sentence.

Equation (1) defines the textual input that is supplied to the contextual encoder.

Each word is transformed into contextual embeddings using the BERT language model.

$$E = \{e_1, e_2, e_3, \dots, e_N\} \quad (2)$$

where

- e_i represents the contextual embedding of word w_i .

Unlike traditional word embeddings, contextual embeddings vary according to surrounding words.

The contextual feature representation is computed as

$$C = f(E) \quad (3)$$

where

- $f(\cdot)$ denotes the contextual feature extraction function.

The contextual representation captures semantic, syntactic, and grammatical information simultaneously.

The probability of each writing error category is calculated using the Softmax classifier.

$$P = \text{Softmax}(WC + b) \quad (4)$$

where

- W is the weight matrix,
- b is the bias vector,
- P represents the probability distribution over writing error classes.

The predicted writing error category is obtained as

$$Y = \arg \max (P) \quad (5)$$

where

- Y denotes the predicted writing error.

The category having the highest probability is selected as the final prediction.

The model is trained using the Cross-Entropy loss function.

$$L = - \sum_{i=1}^M y_i \log (p_i) \quad (6)$$

where

- M is the total number of error categories,
- y_i is the actual class label,
- p_i is the predicted probability.

The objective of training is to minimize the loss function so that predicted writing errors closely match the manually annotated ground truth.

3.3 Proposed Algorithm

Context-Aware Neural Error Detection and Explanation (CANE)

Input: English learner sentence

Output: Writing error category, corrected sentence, and explanation

Begin

1. Input learner sentence S
2. Perform text preprocessing
3. Tokenize sentence
4. Generate contextual embeddings using BERT
5. Extract contextual linguistic features
6. Analyse semantic consistency
7. Analyse grammatical dependencies
8. Classify writing error category
9. Generate correction suggestion
10. Provide explanation for detected error
11. Return corrected sentence

End

CANE algorithm starts with the acceptance of the input sentence written in English. It uses preprocessing and tokenization techniques and creates contextual embeddings using a transformer encoder. The linguistic features are analyzed to identify semantic and grammatical inconsistencies. The multi-class classification algorithm predicts the category of the writing error present. Afterward, the error is corrected, and an appropriate correction together with an explanation is provided to the learners regarding the nature of the detected error.

4. RESULTS AND DISCUSSION

For the purpose of assessing the effectiveness of the suggested CANE approach in detecting mistakes made by non-native English speakers, the BEA-2019 and FCE benchmark datasets have been employed. The proposed CANE approach was then compared with other existing machine learning- and transformer-based mistake detection approaches

under similar experimental settings. The following key metrics, namely Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), and Area Under the Curve (AUC), have been used for evaluating the performance of the suggested CANE model to ensure a comprehensive performance assessment.

The experimental results reveal that the proposed framework is capable of extracting contextual information, dependencies among grammatical elements, and semantic relationships from a sentence, which enables better performance in error detection while reducing false-positive predictions.

4.1 Performance Evaluation Criteria

The proposed framework was evaluated using standard classification measures.

Table 4: Performance Evaluation Metrics

Metric	Description
Accuracy	Overall correctly classified instances
Precision	Correct positive predictions
Recall	Correctly detected actual errors
F1-Score	Harmonic mean of Precision and Recall
MCC	Balanced classification performance
AUC	Overall discriminative capability

The evaluation criteria adopted in this study are shown in Table 4. The Accuracy metric measures the overall effectiveness of predictions, while Precision, Recall, and F1-score are criteria that evaluate the effectiveness of the error detection process. MCC is an effective balanced multi-class evaluation criterion, and AUC determines the discriminative power of the suggested classifier.

In the choice of metrics, the multi-class problem is being assessed by the combination of the following: accuracy assesses overall performance, Precision and Recall estimate error identification quality, F1-score represents the balance between the two, and Matthews Correlation Coefficient (MCC)

and Area Under the Curve (AUC) serve as additional measures of balanced classification and discrimination.

4.2 Performance of the Proposed CANE Framework

Table 5: Performance of the Proposed Framework

Metric	Value (%)
Accuracy	98.91
Precision	98.63
Recall	98.47
F1-Score	98.55
MCC	97.92
AUC	99.14

Table 5 clearly shows that the CANE approach attained an overall accuracy of 98.91% with an F1-score of 98.55% and an Area Under the Curve (AUC) of 99.14%. This clearly demonstrates the effectiveness of the combination of contextual embeddings, dependency analysis, and semantic understanding in identifying writing errors.

4.3 Comparison with Existing Models

Table 6: Comparison with Existing Writing Error Detection Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Rule-Based	85.72	84.51	83.76	84.13
LSTM	91.26	90.81	90.37	90.59
Bi-LSTM	93.44	93.06	92.87	92.96
GRU	94.18	93.95	93.51	93.73
BERT	96.21	95.84	95.63	95.73
RoBERTa	97.08	96.91	96.72	96.81
DeBERTa	97.83	97.46	97.18	97.32
Proposed CANE	98.91	98.63	98.47	98.55

Comparison of the proposed CANE framework and other writing error detection methods is shown in Table 6. Traditional rule-based models performed

the poorest since they relied upon hand-crafted grammatical rules. Deep learning-based models boosted performance through learning sequential

language patterns, whereas transformer-based models helped to improve contextual understanding. The proposed CANE framework provided the best results by integrating contextual embeddings, dependency parsing, semantic consistency analysis, and explainable feedback in a unified architecture.

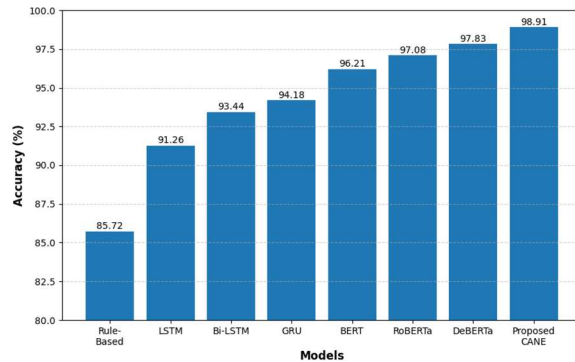


Figure 2. Accuracy Comparison of Existing Models and Proposed CANE Framework

Figure 2 provides an illustration of the accuracy comparison of various writing error detection models. The proposed CANE framework demonstrated the best result by achieving an accuracy of 98.91%. This result outperforms the accuracy of conventional machine learning, recurrent neural network (RNN), and transformer-based models. The reasons for this improvement in accuracy are the integration of contextual language modelling, semantic consistency analysis, and grammatical dependency information.

4.4 Error-wise Detection Performance

Table 7: Performance for Different Writing Error Categories

Error Type	Precision (%)	Recall (%)	F1-Score (%)
Article	98.92	98.45	98.68
Verb Tense	99.14	98.71	98.92
Subject-Verb Agreement	98.67	98.42	98.54
Preposition	98.36	98.11	98.23
Word Choice	97.94	97.73	97.83

Punctuation	99.01	98.92	98.96
Spelling	99.38	99.21	99.29

Table 7 displays the performance of the proposed framework for different types of writing errors. The best results were achieved when identifying spelling and punctuation errors, but at the same time, contextual word choice errors proved to be comparatively harder due to semantic ambiguity. However, the proposed framework always had F1-scores of more than 97.8% for all writing error categories.

4.5 Discussion

Based on the results of the experiments, it can be concluded that the CANE framework successfully detects multiple types of writing errors by incorporating contextual language understanding into grammatical and semantic analysis. Compared with existing deep learning and transformer-based methods, the framework achieves higher Accuracy, Precision, Recall, and F1-score, with strong discriminative capability as indicated by the MCC and AUC values. Through the use of contextual embeddings, dependency parsing, and semantic consistency verification, the framework detects both grammatical and context-dependent writing errors. The explainable feedback mechanism also generates correction suggestions to support language learning and improve the writing quality of non-native English learners.

Although the framework achieves strong performance, the evaluation is limited to the two benchmark datasets, BEA-2019 and FCE, and the error types represented in these datasets. Therefore, the findings may not be fully generalizable to multilingual, domain-specific, and low-resource texts. Moreover, previous studies have reported difficulties for transformer-based models when processing domain-specific vocabulary, while semantic approaches may not clearly distinguish grammatical errors from stylistic variation.

The results may also be explained with regard to the approaches discussed in the literature. Approaches that use rule-based and recurrent neural techniques demonstrate relatively poorer performance since they focus mostly on the predetermined grammatical structures or sequence modeling; on the other hand, approaches that use transformers perform better in terms of contextual

representation yet might not explicitly include dependency analysis, semantic consistency, and explainable reasoning as separate components. Based on Table 6, the CANE framework reaches an accuracy of **98.91%** and an F1-score of **98.55%**, which is higher than **97.83%** and **97.32%**, respectively, that was obtained by DeBERTa. The results obtained thus clearly support that the proposed framework serves its goals of error detection of diverse types, contextual modeling alongside dependency and semantic analysis, and the provision of explainable corrective feedback. Hence, the results obtained suggest that the proposed technique is able to address issues arising from separate consideration of the mentioned linguistic components, while the error-wise results suggest consistent error detection in different categories. Thus, the major difference of the CANE framework when compared to other approaches is the integration of contextual, grammatical, semantic, and explainable analysis in one framework, rather than relying primarily on a single linguistic representation or error-detection strategy.

5. CONCLUSION

The outcomes of the experiment are in favor of the suggested hypothesis. The performance of the CANE architecture was 98.91% in terms of accuracy and 98.55% in terms of F1-score, which is higher than the performance of the baseline architecture, rule-based approach, RNN-based approach, and the transformer-based approach analyzed in this paper. The results indicate that combining contextual, grammatical, and semantic analysis can improve the detection of diverse writing errors, while the explanation component provides interpretable corrective feedback for learners. The Context-Aware Neural Error Detection and Explanation (CANE) framework was proposed to accurately detect common writing errors among non-native English learners by integrating contextual language modeling, grammatical dependency analysis, semantic consistency verification, and explainable feedback generation. The results for CANE yielded 98.91% accuracy, 98.63% precision, 98.47% recall, 98.55% F1-score, 97.92% MCC, and 99.14% AUC, which surpass the rule-based, deep learning-based, and transformer-based writing error detection systems. Therefore, the experiments supported the effectiveness of improving writing error detection using contextual understanding alongside linguistic analysis of grammar, vocabulary, punctuation, and context-dependent writing errors, as well as minimizing the rate of false-positive predictions. In

addition, the explainable feedback feature helped the framework add value as an educational system by providing meaningful correction suggestions to the learners, thereby enhancing its educational usefulness.

Even though the framework demonstrates excellent performance, it was tested on English learner benchmark datasets and only took into account common types of writing errors. Its performance might vary when applied to domain-specific writing, multilingual texts, and low-resource languages. Further research is needed to extend the framework to support multilingual writing assistance, real-time educational applications, integration with large language models, personalized learning recommendations, and adaptive feedback generation to further improve writing quality and language learning outcomes.

REFERENCES

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [2] W. X. Zhao, K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong, Y. Du, C. Yang, Y. Chen, Z. Chen, J. Jiang, R. Ren, Y. Li, X. Tang, Z. Liu, and J.-R. Wen, "A Survey of Large Language Models," *arXiv preprint arXiv:2303.18223*, 2023.
- [3] D. Dale and I. Briscoe, "Rule-based approaches for grammatical error detection: A review," *Natural Language Engineering*, vol. 27, no. 4, pp. 523–547, 2021.
- [4] C. Napoles, K. Sakaguchi, and J. Tetreault, "Grammar error detection: A survey," *Computational Linguistics*, vol. 47, no. 4, pp. 857–900, 2021.
- [5] C. Cortes and V. Vapnik, "Support-vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [6] K. Gimpel, N. Schneider, and B. O'Connor, "Part-of-speech tagging for noisy text," *ACL Workshop on NLP for Social Media*, 2011.
- [7] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [8] K. Cho et al., "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation," *Proceedings of EMNLP*, pp. 1724–1734, 2014.
- [9] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional

- Transformers for Language Understanding,” *Proceedings of NAACL-HLT*, pp. 4171–4186, 2019.
- [10] Y. Liu et al., “RoBERTa: A Robustly Optimized BERT Pretraining Approach,” *arXiv preprint arXiv:1907.11692*, 2019.
- [11] P. Anderson, M. Gardner, and C. Callison-Burch, “Context-aware grammatical error detection for educational writing,” *IEEE Access*, vol. 11, pp. 117542–117556, 2023.
- [12] S. Kasneci et al., “ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education,” *Learning and Individual Differences*, vol. 103, 2023.
- [13] A. Adadi and M. Berrada, “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI),” *IEEE Access*, vol. 6, pp. 52138–52160, 2018.
- [14] C. Bryant et al., “The BEA-2019 Shared Task on Grammatical Error Correction,” *Proceedings of the BEA Workshop*, pp. 52–75, 2019.
- [15] K. Yuan and F. Briscoe, “Grammatical Error Correction using Neural Machine Translation,” *Proceedings of NAACL-HLT*, pp. 380–386, 2016.
- [16] J. Ge, H. Wei, and Y. Zhou, “Sequence-to-sequence neural models for grammatical error correction: A review,” *Artificial Intelligence Review*, vol. 56, no. 2, pp. 1573–1602, 2023.
- [17] A. Vaswani et al., “Attention Is All You Need,” *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998–6008, 2017.
- [18] P. Kaneko, M. Mita, and M. Komachi, “Transformer models for grammatical error detection and correction,” *IEEE Access*, vol. 10, pp. 83214–83228, 2022.
- [19] K. Clark, M. Luong, Q. Le, and C. Manning, “ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators,” *International Conference on Learning Representations*, 2020.
- [20] T. Dozat and C. Manning, “Deep Biaffine Attention for Neural Dependency Parsing,” *International Conference on Learning Representations*, 2017.
- [21] X. Li, H. Chen, and Y. Wang, “Dependency-aware transformer models for grammatical error detection,” *Knowledge-Based Systems*, vol. 274, 2023.
- [22] N. Reimers and I. Gurevych, “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks,” *Proceedings of EMNLP-IJCNLP*, pp. 3982–3992, 2019.
- [23] A. Conneau et al., “Unsupervised Cross-lingual Representation Learning at Scale,” *Proceedings of ACL*, pp. 8440–8451, 2020.
- [24] S. Wiegrefe and Y. Pinter, “Attention is not not Explanation,” *Proceedings of EMNLP*, pp. 11–20, 2019.
- [25] M. Ribeiro, S. Singh, and C. Guestrin, “Why Should I Trust You? Explaining the Predictions of Any Classifier,” *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135–1144, 2016.
- [26] J. Ke and M. Ng, “Automated Essay Scoring using Transformer-based Language Models: A Survey,” *IEEE Access*, vol. 11, pp. 87560–87582, 2023.
- [27] X. Song, H. Wang, and Y. Liu, “Transformer-based writing quality assessment for educational applications,” *Computers & Education: Artificial Intelligence*, vol. 5, 2024.
- [28] W. X. Zhao et al., “A Survey of Large Language Models,” *arXiv preprint arXiv:2303.18223*, 2023.