Rev. Roum. Sci. Techn.– Électrotechn. et Énerg.
Vol. **70**, *3*, pp. 391–396, Bucarest, 2025

**Automatique et ordinateurs**
**Automatics and computers**

# ENHANCING CONVERSATIONAL AGENTS USING ROTATIONAL ATTENTION AND GATED SPLINE MODULES

MUTHUKUMARAN NARAYANAPERUMAL[1], VIGNESH ARUMUGAM[2]

**Keywords: Natural language understanding; Conversational AI; Enhanced T5 Model; Contextual dual-axis rotational attention; Neural-spline gated linear units.**

**In natural language understanding, transformer models like T5 and GPT have achieved strong results in generating contextually relevant responses. However, limitations such as static self-attention in T5 and unidirectional context in GPT hinder their ability to capture deeper inter-token dependencies and nuanced semantics. To address these challenges, we propose an enhanced T5 (ET5) architecture integrating two novel modules: contextual dual-axis rotational attention (CDARA) and neural-spline gated linear units (NS-GLU). CDARA facilitates attention across both token and feature dimensions, while NS-GLU introduces adaptive spline-activated gating for improved nonlinear representation. Experiments on NarrativeQA, SQuAD, MultiWOZ, and DailyDialog show that ET5 consistently outperforms PEGASUS, GPT-3, and T5-LSTM FusionNet. ET5 achieves superior BERTScore (up to 0.971), BLEU (up to 0.77), and lower word error rate (WER) (as low as 0.13), confirming its effectiveness in generating fluent, accurate, and semantically rich responses. These results position ET5 as a promising advancement in transformer-based conversational AI systems.**

## 1. INTRODUCTION

Transformer-based architecture has become the backbone of modern conversational agents, with models such as generative pre-trained transformer (GPT-3) and text-to-text transfer transformer (T5) demonstrating remarkable capabilities in generating fluent, contextually relevant, and diverse responses. Despite their success, both models exhibit fundamental limitations that restrict their performance in complex dialogue generation and question-answering tasks.

T5, while offering a flexible text-to-text framework with an encoder-decoder structure, is hindered by two critical issues. First, its standard self-attention mechanism operates primarily in a unidirectional manner within each encoder and decoder layer, limiting its ability to capture multi-scale contextual dependencies and feature interactions simultaneously. Second, the position-wise feed-forward network (FFN) used after attention layers processes each token independently with fixed activation dynamics, failing to adaptively modulate information based on global sequence structure or token-specific relevance. This leads to suboptimal handling of nuanced semantics, especially in multi-turn conversations or when contextual reasoning is required.

On the other hand, GPT-3, with its autoregressive decoder-only design, generates tokens based on a left-to-right context window. While powerful in fluent text generation, GPT-3 cannot perform bidirectional context encoding, which is crucial for tasks that require understanding both preceding and succeeding content, such as question answering or follow-up reasoning in dialogue. Additionally, GPT-3 suffers from over-reliance on surface-level patterns and often produces generic, verbose, or factually inconsistent responses, especially when encountering ambiguous prompts or low-resource domains. Its sheer size, although advantageous for generalization, introduces inefficiencies in fine-tuning and adaptation, making it impractical for personalized or domain-specific applications.

These observed limitations motivate the need for architectural enhancements that can: (i) capture deep contextual relationships across both time and feature dimensions, (ii) offer adaptive non-linear transformations beyond rigid FFN behavior, and (iii) maintain a bidirectional and generative framework suited for open-domain response tasks.

To address these challenges, we propose an Enhanced T5 model (ET5) that integrates two novel components: (i) Contextual Dual-Axis Rotational Attention (CDARA), a dual-dimensional attention mechanism that captures rich semantic dependencies across token positions and embedding features; and (ii) Neural-Spline Gated Linear Units (NS-GLU), which replace the feed-forward network with a spline-based adaptive gating module, enabling context-sensitive non-linear feature modulation. These enhancements are designed to address the inherent inflexibility of standard transformer architectures and enhance the system's capability to produce responses that are more fluent, contextually relevant, and accurate.

Empirical evaluations on multiple question-answering and dialogue datasets (NarrativeQA, SQuAD, MultiWOZ, and DailyDialog) demonstrate that ET5 consistently outperforms baseline architectures in bidirectional encoder representations from transformers score (BERTScore), bilingual evaluation understudy (BLEU), and word error rate (WER) metrics. The findings suggest that enriching attention and feed-forward components within transformer frameworks offers a promising direction for advancing the state of conversational AI.

## 2. RELATED WORKS

Recent advances in artificial intelligence (AI) have led to significant improvements in question answering (QA) chatbots, driven by innovations in datasets, text representation, and model architectures [1–6]. Early systems primarily relied on handcrafted features and statistical models, but the emergence of large-scale datasets such as NarrativeQA and SQuAD [7–10], as well as dialogue-specific corpora like the Cornell Movie Dialogues [11,12], has facilitated data-driven training approaches for open-domain conversation systems.

Word representation has evolved from traditional frequency-based techniques to dense vector embeddings [13,14]. For instance, term frequency–inverse document frequency (TF-IDF) has been used in early retrieval-based models [15–18]. At the same time, more recent efforts incorporate distributed word representations such as GloVe or domain-adaptive sparse embeddings, offering improved semantic understanding for neural architectures.

In terms of modeling, a broad spectrum of machine learning

[1,2] Sri Eshwar College of Engineering, Coimbatore – 641202, Tamil Nadu, India.
E-mails: muthukumaran.n@sece.ac.in, vignesh.a@sece.ac.in

DOI: 10.59277/RRST-EE.2025.70.3.18

(ML) and deep learning (DL) strategies has been explored. Classical ML models [19,20], alongside deep learning architectures, have enabled more fluent and context-aware chatbot responses. The sequence-to-sequence (Seq2Seq) model has become foundational for response generation and has been extended through attention mechanisms to improve output fluency and contextual relevance.

Recently, transformer architectures have revolutionized chatbot development, primarily due to their scalability and exceptional capability in capturing long-distance contextual relationships [21]. Prior research has validated the efficiency of models like BERT, GPT, and T5 in enabling comprehensive end-to-end chatbot training.

### 2.1 LITERATURE EMPLOYING THE TRANSFORMER MODEL FOR RESPONSE GENERATION

A transformer-based model, ALSI-transformer, was proposed to improve automatic code comment generation by leveraging both lexical and syntactic features through a novel code-aligned type (CAT) sequence [15]. This alignment enhances semantic representation by combining code tokens with syntax structures. The model outperforms existing baselines on Java datasets and reduces training time. However, it lacks support for project-level comment generation and cross-function context understanding. Its reliance on language-specific parsers also limits generalizability to other programming languages, suggesting a need for future enhancements across diverse codebases.

To enhance abstractive summarization, PEGASUS employs a unique pre-training strategy known as gap-sentence generation, in which essential sentences are masked and inferred using the remaining text [16]. This method enables the model to capture representations relevant to summarization, delivering top-tier results across 12 benchmark datasets. It is also effective in environments with limited data. However, PEGASUS relies heavily on heuristic-based sentence selection, which may not generalize well across all domains. Additionally, its computational requirements during pre-training are high, posing scalability concerns for broader real-world applications.

To enhance summarization in the psychological domain, a hybrid model combining T5 and LSTM, known as T5 LSTM FusionNet, was introduced for improved contextual and sequential representation [17]. The model demonstrated improved accuracy, precision, and ROUGE scores compared to standalone and transformer-based models. It effectively captured unigram and bigram overlaps, contributing to more coherent summaries. However, the fusion mechanism remains suboptimal, particularly in preserving long-range dependencies and sequence flow. Moreover, the model's adaptability across multilingual datasets and clinical decision-making contexts remains to be validated, indicating potential areas for further improvement.

A novel text-based steganography method was introduced using GPT-3 to generate natural language stego text for covert communication over software-defined radios (SDRs) [18]. The approach encodes secret bits into abbreviations and uses GPT-3 to generate plausible text expansions, achieving low perplexity (2.75), high human opinion scores (4.052), and high decoding accuracy. It also demonstrated resistance to detection, with low recognition accuracy (0.5245). Despite its effectiveness, the method relies on a fixed language model and pre-agreed encoding schemes. Future improvements could explore customized language models and adaptive text frameworks for enhanced security and versatility.

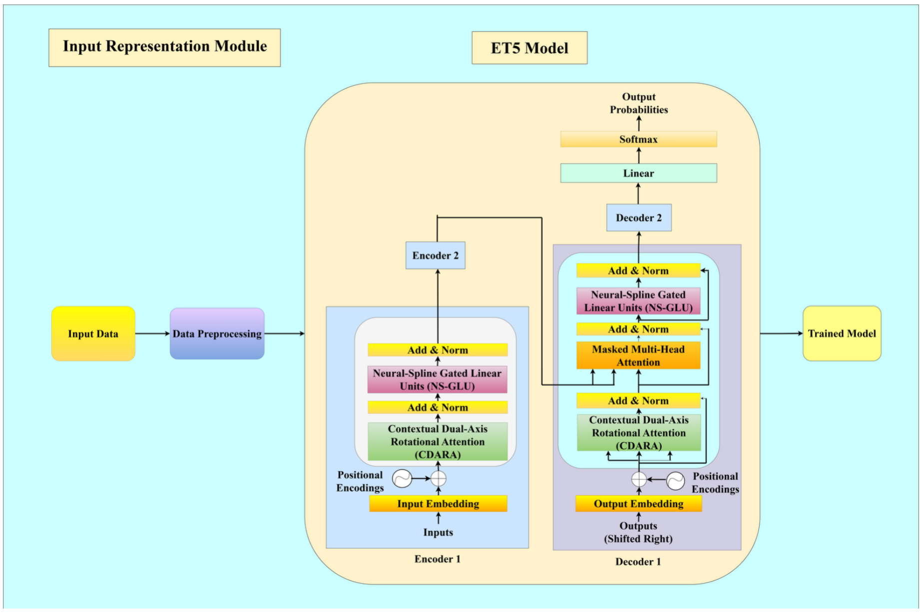


Fig. 1. – Architecture of ET5 Model.

## 3. **PROPOSED METHODOLOGY**

This section describes the architectural design and internal components of the proposed Enhanced T5 (ET5) model. It outlines the input processing workflow, encoder-decoder mechanisms, and the integration of two novel modules: CDARA and NS-GLU to improve semantic representation and response generation.

### 3.1 FUNCTIONAL LAYOUT OF THE ET5 ARCHITECTURE

The Enhanced T5 (ET5) model advances the original T5 framework by addressing its limitations in semantic understanding and adaptive feature learning. ET5 introduces two core innovations: CDARA and NS-GLU, which are embedded within both the encoder and decoder stacks. The model begins with preprocessing and tokenization, followed by input embedding and positional encoding. CDARA captures rich contextual relationships by applying dual-stream attention over both token (temporal) and feature (channel) dimensions, enhanced through rotational transformations. This facilitates a deeper understanding of semantic and syntactic dependencies. NS-GLU replaces the standard feed-forward network with a spline-activated gated mechanism, enabling smooth, learnable non-linear modulation of features. In the decoder, masked multi-head attention supports autoregressive generation by processing shifted output tokens, after which CDARA and NS-GLU refine context and feature integration again. Final outputs are projected to the vocabulary space via a linear layer and a softmax function. Through this architecture, ET5 improves upon T5 and GPT variants, achieving higher performance across metrics such as BLEU, BERTScore, and WER on benchmarks including NarrativeQA, SQuAD, MultiWOZ, and DailyDialog, demonstrating superior generalization, semantic fidelity, and response quality in natural language tasks. Figure 1 represents the architecture of ET5 model.

## 3.2 COMPONENTS OF THE ET5 MODEL

### 3.2.1 INPUT REPRESENTATION AND POSITIONAL ENCODING

The input representation and positional encoding stage initiates the ET5 model by converting raw text into structured semantic embeddings. The input text is first preprocessed through tokenization, normalization, and formatting to ensure compatibility with the transformer. These tokens are then embedded into dense vectors via an input embedding layer that captures essential lexical semantics. To preserve sequence order, fixed sinusoidal positional encodings are added, providing parameter-free positional awareness without additional training overhead. This enables the model to recognize relative token positions effectively. During decoding, output tokens are right-shifted and embedded in a similar manner, with fixed positional encodings applied to maintain sequence alignment for autoregressive generation. This phase ensures that both encoder and decoder components receive inputs that are contextually enriched and position-aware, laying a strong foundation for accurate and coherent text generation in downstream tasks.

### 3.2.2 ENCODER DESIGN

The encoder in the Enhanced T5 (ET5) model is designed to capture deep contextual dependencies from the input sequence using a stack of modified transformer layers. Each encoder block replaces traditional self-attention with CDARA to model both token-level and feature-level interactions. This is followed by NS-GLU, which substitute standard feed-forward layers to introduce flexible, non-linear feature transformations. Residual connections and layer normalization are applied throughout to stabilize learning and preserve gradient flow.

### 3.2.3 DECODER DESIGN

The decoder in the Enhanced T5 (ET5) model is designed to generate responses in an auto-regressive manner by integrating masked multi-head attention, CDARA, and NS-GLU. Initially, masked attention enables the decoder to attend only to previous tokens, preserving the causality required for sequential decoding. This is followed by encoder-decoder attention using CDARA, which enhances positional alignment and context flow by capturing cross-dimensional interactions between encoder outputs and decoder states. Finally, NS-GLU fuses the attended features with learnable, spline-based gating, enabling adaptive control over the information flow and enhancing the decoder's expressive power. Together, these components ensure the generation of accurate, fluent, and context-aware responses.

### 3.2.4 CDARA

CDARA is an innovative attention framework designed to enhance the model's ability to recognize both temporal and feature-level dependencies within the input data. Unlike traditional self-attention, which focuses solely on token-wise relationships, CDARA introduces a dual-axis approach by performing attention across both the temporal (token) and feature (channel) dimensions. A rotational transformation is interleaved between these axes to facilitate richer cross-dimensional interactions. This is followed by a gated fusion mechanism that adaptively combines the outputs from both attention paths, enabling dynamic weighting of contextual signals. The result is a semantically enriched representation that balances inter-token coherence with intra-feature expressiveness. Figure 2 represents the architecture of the CDARA module.

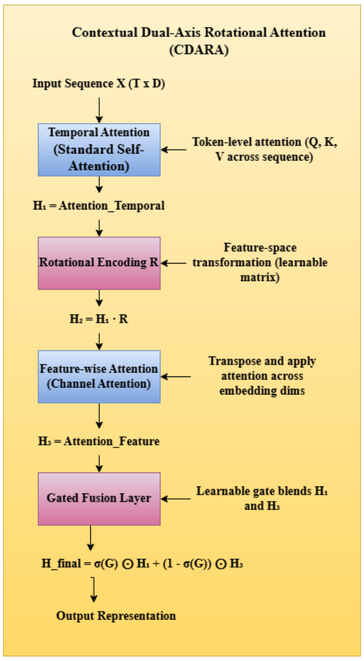


Fig. 2. – CDARA Architecture.

#### 3.2.4.1 INPUT SEQUENCE

Let the input sequence be:

$$X \in R^{T*D}. \tag{1}$$

Equation (1) represents the input matrix with $T$ tokens and $D$ hidden dimensions. This input is passed to the Temporal Attention block

#### 3.2.4.2 TEMPORAL ATTENTION (TOKEN-WISE ATTENTION)

Query, Key, Value projections:

$$Q = XW_Q, K = XW_K, V = XW_V. \tag{2}$$

Equation (2) defines the linear transformations used to obtain query, key, and value matrices for temporal attention. Scaled Dot-Product Attention:

$$Attn_{temporal}(Q, K, V) = softmax\left(\frac{Qk^T}{\sqrt{d_k}}\right)v. \quad (3)$$

Equation (3) computes the attention-weighted values across the token dimension.

Output of temporal attention:

$$H_1 = Attn_{temporal}(X \in R^{T*D}). \quad (4)$$

Equation (4) represents the context-aware output after applying token-wise attention.

#### 3.2.4.3 ROTATIONAL ENCODING (FEATURE MIKING LAYER)

Rotational transformation:

$$H_2 = H_1.R, \quad R \in R^{D*D}. \quad (5)$$

Equation (5) applies a learnable rotation matrix R to encourage feature-level mixing within the embedding space.

#### 3.2.4.4 FEATURE-WISE ATTENTION (CHANNEL-WISE ATTENTION)

Transpose and project:

$$Q_f = H_2 w_f^Q, k_f = H_2^T w_f^k, v_f = H_2^T w_f^v. \quad (6)$$

Equation (6) projects the transposed feature into queries, keys, and values for channel-wise attention.

Feature-wise attention:

$$Attn_{feature}(Q_f, K_f, V_f) = softmax\left(\frac{Q_f k_f^T}{\sqrt{d_k}}\right)V_f. \quad (7)$$

Equation (7) computes scaled dot-product attention over the feature axis (per token).

Output after feature attention:

$$H_3 = (Attn_{feature}(H_2^T))^T \in R^{T*D}. \quad (8)$$

Equation (8) denotes the feature-enhanced presentation transposed back to match the original shape.

#### 3.2.4.5 GATED FUSION LAYER (CONTEXT FUSION)

Learnable gate:

$$G = \left(H_1 w_g + H_3 w_9^1\right) \in R^{T*D}. \quad (9)$$

Equation (9) defines the gating vector using a sigmoid function over a linear combination of temporal and feature paths.

Final fused output:

$$H_{final} = G \odot H_1 + (1 - G) \odot H_3. \quad (10)$$

Equation (10) performs element-wise fusion, balancing both attention sources dynamically.

#### 3.2.4.6 OUTPUT OF CDARA

$$H_{cDARA} = H_{final} \in R^{T*D}. \quad (11)$$

Equation (11) provides the final output of the CDARA block, ready to be passed to the next module.

### 3.2.5 NS-GLU

NS-GLU is a novel replacement for the conventional feed-forward network in transformer architectures. It introduces a dual-path computation mechanism: a primary linear transformation path and a gating path modulated by a learnable spline activation function. Unlike traditional activation functions, the spline activation enables smooth, flexible, and shape-controllable non-linearities. The gated output adaptively scales the primary features, enhancing expressiveness and dynamic range. This structure allows for the model to learn complex transformations more efficiently, thereby improving its representational power with minimal parameter overhead. Figure 3 represents the architecture of the NS-GLU module.

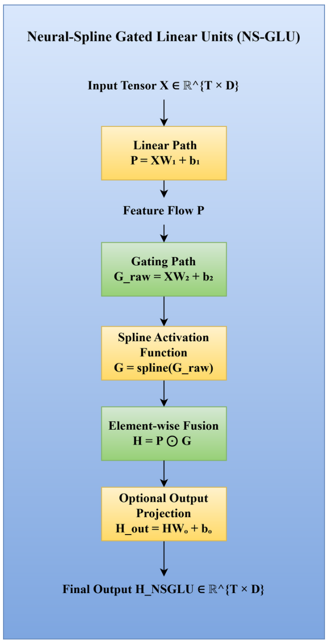


Fig. 3 – NS-GLU Architecture.

### 3.2.6 OUTPUT GENERATION LAYER

The decoder output is linearly transformed into logits, passed through a softmax function to compute token probabilities, and the highest-probability token is selected for autoregressive generation.

## 4. RESULT AND DISCUSSION

### 4.1. EXPERIMENTAL CONFIGURATION

The proposed model was developed using Python 3.10 and TensorFlow 2.11.0 within the Spyder IDE environment. Training and evaluation were conducted using benchmark datasets: NarrativeQA, SQuAD, MultiWOZ, and DailyDialog. Before training, datasets underwent standard preprocessing, including tokenization, lowercasing, and padding to ensure uniformity across input sequences. The ET5 model was trained for 100 epochs with a batch size of 64 and a learning rate of 0.001. The model's hidden dimensionality was set to 1024, and each CDARA module employed 8 attention heads to capture temporal and feature-wise dependencies. A dropout rate of 0.3 mitigated overfitting, while the epsilon value for layer normalization was fixed at 1e-6 for numerical stability. The Adam optimizer was used to accelerate convergence due to its adaptive learning rate and momentum-based updates. A 10% validation split was applied, with early stopping enabled to avoid overfitting. Model training and testing were performed on a medium-scale system featuring an NVIDIA GTX 1660 Ti GPU (6 GB VRAM), 32 GB RAM, and an Intel Core i7-10700 processor. Training durations ranged between 70 and 90 hours, depending on the dataset. Hyperparameters were tuned empirically using grid search to ensure optimal model performance.

## 4.2. RESULT ANALYSIS

The ET5 framework integrates advanced components—Contextual Dual-Axis Rotational Attention (CDARA) and a Bidirectional Neural-Spline Gated Linear Unit (NS-GLU)—to enhance contextual understanding in dialogue generation. Training and evaluation were conducted on four benchmark datasets: NarrativeQA, SQuAD, MultiWOZ, and DailyDialog. The best-performing results from each dataset variant were compared against several state-of-the-art models, including ALSI-Transformer, PEGASUS, T5-LSTM FusionNet, and GPT-3, to assess the effectiveness of the proposed architecture.

Table 1 presents performance comparisons using BERTScore, BLEU, and Word Error Rate (WER). ET5 consistently surpasses all baselines in BERTScore and BLEU, demonstrating its ability to generate semantically rich and fluent responses. It also achieves the lowest WER, indicating superior accuracy in sequence generation. These findings validate the impact of ET5's architectural enhancements—CDARA and NS-GLU—in improving contextual representation and overall response quality in transformer-based conversational systems.

*Table 1*

Evaluation results of the ET5 model vs. state-of-the-art models.

| Metrics | Model | NarrativeQA | SQuAD | MultiWOZ | DailyDialog |
|---|---|---|---|---|---|
| BERTScore | ALSI-Transformer | 0.921 | 0.927 | 0.902 | 0.896 |
| | PEGASUS | 0.928 | 0.934 | 0.911 | 0.903 |
| | T5-LSTM FusionNet | 0.932 | 0.938 | 0.916 | 0.908 |
| | GPT-3 | 0.949 | 0.953 | 0.932 | 0.923 |
| | ET5 (proposed) | 0.968 | 0.971 | 0.953 | 0.946 |
| BLEU | ALSI-Transformer | 0.692 | 0.721 | 0.701 | 0.678 |
| | PEGASUS | 0.705 | 0.739 | 0.717 | 0.689 |
| | T5-LSTM FusionNet | 0.711 | 0.746 | 0.723 | 0.694 |
| | GPT-3 | 0.738 | 0.763 | 0.742 | 0.713 |
| | ET5 (proposed) | 0.753 | 0.778 | 0.765 | 0.739 |
| WER | ALSI-Transformer | 0.214 | 0.196 | 0.231 | 0.245 |
| | PEGASUS | 0.198 | 0.182 | 0.219 | 0.233 |
| | T5-LSTM FusionNet | 0.191 | 0.175 | 0.213 | 0.226 |
| | GPT-3 | 0.171 | 0.157 | 0.194 | 0.209 |
| | ET5 (proposed) | 0.154 | 0.132 | 0.172 | 0.184 |

### 4.2.1. BERTSCORE

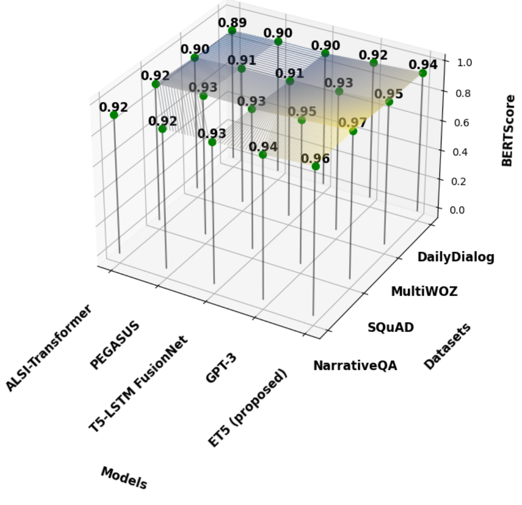


Fig. 4 – BERTScore.

Figure 4 shows 3D BERTScore comparisons, where ET5 outperforms ALSI-Transformer, PEGASUS, and GPT-3, achieving 0.95 on NarrativeQA and 0.94 on DailyDialog, demonstrating improved semantics via CDARA and NS-GLU modules.

Figure 5 presents BLEU score comparisons, where ET5 outperforms ALSI-Transformer, PEGASUS, and GPT-3, achieving 0.77 on MultiWOZ and 0.76 on DailyDialog and SQuAD, showing superior fluency and relevance in generation tasks.

### 4.2.2. BLEU SCORES

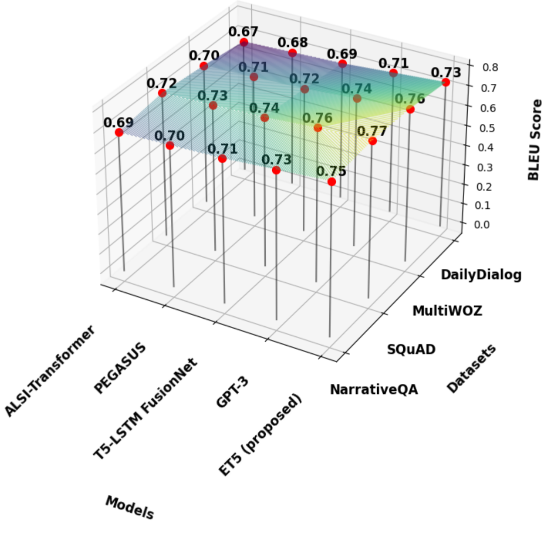


Fig. 5 — BLEU Score.

### 4.2.3. WER SCORES

Figure 6 shows the Word Error Rate (WER) across models and datasets. The proposed ET5 achieves the lowest WER, with 0.13 on NarrativeQA and 0.15 on SQuAD, outperforming GPT3 and PEGASUS, indicating higher accuracy and fewer generation errors.

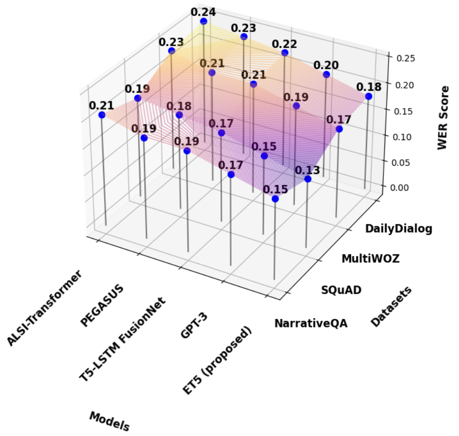


Fig. 6 – WER Score.

## 5. CONCLUSION AND FUTURE SCOPE

This research presents the Enhanced T5 (ET5) architecture, a novel extension of the traditional T5 model designed to overcome the inherent limitations in contextual understanding and feature modulation. By incorporating CDARA and NS-GLU, ET5 significantly enhances the ability to capture intricate dependencies across both token sequences and embedding dimensions. CDARA enables comprehensive dual-dimensional attention, while NS-GLU introduces a smooth, adaptive non-linearity to replace static feed-forward layers. Empirical evaluation across four benchmark datasets, including NarrativeQA, SQuAD, MultiWOZ, and DailyDialog, demonstrates that ET5 consistently outperforms established models such as GPT - 3, PEGASUS, and T5 LSTM FusionNet, achieving BERTScore scores of up to 0.971, BLEU scores of 0.77, and WER as low as 0.13. The model's strong performance confirms its effectiveness in producing more accurate, semantically rich, and contextually grounded responses.

While ET5 shows notable improvements, future work includes incorporating reinforcement learning with human feedback for personalization, extending it to multimodal inputs (text, images, audio), and optimizing for low-resource settings. Emphasis will also be placed on enhancing interpretability and fairness for transparent, responsible AI communication.

## CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Muthukumaran Narayanaperumal: conceptualization, data curation, methodology, supervision, writing – review & editing.
Vignesh Arumugam: conceptualization, data curation, methodology, software, investigation, writing – original draft, visualization.

*Received on 3 November 2024*

## REFERENCES

1. M. Ganga, G. Jasmine, N. Muthukumaran, M. Veluchamy, *Red fox-based fractional order fuzzy PID controller for smart LED driver circuit*, Rev. Roum. Sci. Techn. – Électrotechn. Et Énerg., **68,** pp. 395–400 (2023).
2. R. Ahmad, D. Siemon, U. Gnewuch, S. Robra-Bissantz, *Designing personality-adaptive conversational agents for mental health care,* Information Systems Frontiers, **24**, pp. 923–943 (2022).
3. A. Ramaiah, P. Devi Balasubramanian, A. Appathurai, N. Muthukumaran, *Génie biomédical Biomedical Engineering detection of Parkinson's disease via Clifford gradient-based recurrent neural network using multi-dimensional data*, Rev. Roum. Sci. Techn. -Électrotechn. et Énerg, **69** (2024).
4. J. Balakrishnan, Y.K. Dwivedi, *Conversational commerce: entering the next stage of AI-powered digital assistants*, Ann Oper Res (2021).
5. A. Appathurai, A.S.I. Tinu, N. Muthukumaran, *Meg and Pet images-based brain tumor detection using Kapur's Otsu segmentation and sooty optimized Mobilenet classification*, Rev. Roum. Sci. Techn. – Électrotechn. Et Énerg., **69,** *3,* pp. 363–368 (2024).
6. W. Cai et al., *Bandit algorithms to personalize educational chatbots*, Mach Learn, **110**, *9,* pp. 2389–2418 (2021).
7. T.Y. Chen, Y.C. Chiu, N. Bi, R.T.H. Tsai, *Multi-modal Chatbot in intelligent manufacturing*, IEEE Access (2021).
8. S. Gong, M. Li, J. Feng, Z. Wu, L. Kong, *DiffuSeq: Sequence to sequence text generation with diffusion models*, (2022).
9. L. Grassi, C.T. Recchiuto, A. Sgorbissa, *Knowledge-Grounded dialogue flow management for social robots and conversational agents*, Int J Soc Robot, **14**, *5,* pp. 1273–1293, (2022).
10. H. Honda, M. Hagiwara, Question answering systems with deep learning-based symbolic processing, IEEE Access, **7**, pp. 152368–152378 (2019).
11. C. Hsu, C.C. Chang, *Integrating machine learning and open data into social Chatbot for filtering information rumor*, J Ambient Intell Humaniz Comput, **12**, *1,* pp. 1023–1037, (2021).
12. R.B. Lincy, R. Gayathri, *Optimized convolutional neural network for tamil handwritten character recognition,* Intern J Pattern Recognit Artif Intell, **36**, *11*, (2022).
13. M.M. Mohsan, M.U. Akram, G. Rasool, N.S. Alghamdi, M.A.A. Baqai, M. Abbas, *Vision transformer and language model-based radiology report generation*, IEEE Access, **11,** pp. 1814–1824 (2023).
14. K. Palasundram, N. Mohd Sharef, K.A. Kasmiran, A. Azman, *Enhancements to the sequence-to-sequence-based natural answer generation models*, IEEE Access, **8,** pp. 45738–45752 (2020).
15. Y. Park, A. Park, C. Kim, *ALSI-Transformer: transformer-based code comment generation with aligned lexical and syntactic information*, IEEE Access, **11**, pp. 39037–39047 (2023).
16. J. Zhang, Y. Zhao, M. Saleh, P.J. Liu, *PEGASUS: Pre-training with extracted gap-sentences for abstractive summarization*, arXiv (2019)
17. B. Khan, M. Usman, I. Khan, J. Khan, D. Hussain, Y.H. Gu, *Next-generation text summarization: A T5-LSTM FusionNet hybrid approach for psychological data*, IEEE (2025).
18. M.V. Namitha, G.R. Manjula, M.C. Belavagi*, StegAbb: A cover-generating text steganographic tool using GPT-3 language modeling for covert communication across SDRs*, IEEE Access, **12**, pp. 82057–82067 (2024).
19. D. Mylsamy, A. Appathurai, N. Muthukumaran, S. Kuppusamy, *Mojo-based fuzzy agglomerative clustering algorithm with Ed 2 Mt strategy for large-scale wireless sensor networks*, Rev. Roum. Sci. Techn. -Électrotechn. et Énerg, **69** (2024).
20. W.T. Wang, N. Tan, J.A. Hanson, C.A. Crubaugh, A.K. Hara, *Initial experience with a COVID-19 screening chatbot before radiology appointments*, J Digit Imaging, **35**, *5,* pp. 1303–1307 (2022).
21. M.R. Kumar, R. Sundaram, M. Rengasamy, R. Balakrishnan, *Effective feature extraction method for unconstrained environment: local binary pattern or local ternary pattern*, Rev. Roum. Sci. Techn. – Électrotechn. Et Énerg., **69,** *4*, pp. 443–448 (2024).