

Development of an AI Driven Text Simplification and Analogy Generation Platform Using a Pre-Trained BART Model

Alimi O. Maruf¹, James Richard Henshaw², Oluwaseyi Ezekiel Olorunshola³, Adeniyi Usman Adedayo⁴,
Enem A Theophilus⁵, Adamu-Fika Fatimah⁶

¹²³Department of Computer Science, Faculty of Computing, Airforce Institute of Technology Kaduna,
Nigeria

⁴⁵⁶Department of Cyber Security, Faculty of Computing Airforce Institute of Technology Kaduna,
Nigeria

DOI: <https://doi.org/10.51584/IJRIAS.2025.1010000021>

Received: 25 September 2025; Accepted: 30 September 2025; Published: 29 October 2025

ABSTRACT

Understanding complex information can be a challenge for most learners, especially when it is filled with technical terms, abstract ideas, or specialized language. Education, research, and technical communication often suffer when content is too difficult for the intended audience. Simplifying text can help, but simplification alone does not always create the mental connections needed for deeper understanding. This research proposes and develops an AI-driven platform that combines text simplification and analogy generation to make complex information clearer and more relatable. A pre-trained BART model is used to simplify text while preserving meaning, and a Retrieval-Augmented Generation (RAG) process is applied to generate analogies based on user-selected themes such as sports or classrooms. The system is built with Python for the backend and Flutter for the frontend, offering a user-friendly interface for real-time processing. Evaluation using ROUGE and BERTScore confirmed the system's effectiveness. Summarization achieved a ROUGE-1 score of 0.8315, while text simplification reached a BERTScore F1 of 0.9279, indicating high semantic fidelity. Analogy generation maintained F1 scores above 0.7, demonstrating relevance and conceptual clarity. These results confirm the platform's ability to improve comprehension through high-quality simplification and relatable analogies, making it a practical tool for education and accessible communication across diverse domains.

Keywords: Text Simplification, Analogy Generation, BART Model, Retrieval-Augmented Generation, Natural Language Processing, Semantic Preservation

INTRODUCTION

In today's information-rich society, the ability to comprehend and communicate complex ideas is crucial across various domains, including education, research, and professional settings. However, many individuals encounter difficulties in processing dense, technical, or abstract texts due to linguistic complexity and specialized vocabulary. These challenges can impede accessibility and hinder knowledge transfer, particularly for second-language learners and individuals with lower literacy levels or cognitive impairments (Seidenberg, 2013).

Text simplification has emerged as a promising solution to address these challenges by transforming complex content into more readable forms while preserving its meaning. This can be achieved through lexical simplification, which involves replacing difficult words with simpler alternatives, and syntactic simplification, which restructures complex sentences to improve clarity. Studies have demonstrated that these methods can reduce cognitive load and enhance comprehension for a diverse range of learners (Crossley et al., 2011).

Despite the benefits of simplified language, comprehension in technical and academic fields often requires more than just simplified wording; it necessitates building conceptual bridges between unfamiliar and familiar

knowledge. Analogy generation addresses this need by creating meaningful comparisons that facilitate understanding and promote knowledge transfer. For example, abstract concepts like quantum mechanics or cognitive processes can be compared to familiar scenarios, making them more approachable for general audiences (Richland & Simms, 2015).

Advancements in Natural Language Processing (NLP) have significantly improved text simplification techniques. Transformer-based models, such as BART, have been shown to produce fluent and semantically faithful outputs, enhancing the effectiveness of text simplification processes (Lewis et al., 2020). Furthermore, Retrieval-Augmented Generation (RAG) techniques enable large language models to draw relevant knowledge from external databases, improving contextual accuracy and relevance in analogy generation (Lazaridou et al., 2021).

This project builds upon these advancements by integrating a pre-trained BART model for text simplification with a RAG-based analogy generation pipeline. The resulting platform aims to deliver accessible, personalized, and semantically accurate outputs, thereby supporting comprehension in educational, technical, and accessibility-focused contexts.

Statement Of The Problem

In today's information-rich world, understanding complex subjects is essential for students and professionals alike. However, learners often struggle with dense texts, technical jargon, and abstract ideas, exacerbated by the increasing volume of information (Carr, 2010).

Current tools lack personalization to match individual learning styles, providing simplified text that reduces complexity but fails to address deeper cognitive challenges (Richland et al., 2017). For instance, they may preserve meaning but overlook the need for conceptual connections, leading to incomplete comprehension in technical fields (Martinez & Lee, 2023).

To bridge this gap, an integrated system combining text simplification and analogy generation is required to foster relatable connections and enhance knowledge transfer.

Review Of Related Works

Early research on text simplification focused heavily on rule-based approaches, where linguists manually defined rules for breaking down long sentences and replacing complex words. These systems were predictable and interpretable but lacked scalability and often produced overly formal or awkward outputs. Their limitations became more obvious when dealing with nuanced, real-world text, motivating a shift toward data-driven machine learning solutions (Siddharthan, 2014).

The rise of statistical machine translation (SMT) techniques introduced probabilistic models that treated simplification as a translation task, mapping complex sentences to simpler counterparts. These approaches produced better fluency but struggled with meaning preservation and relied heavily on parallel corpora, which were expensive to build. As a result, they were difficult to apply in low-resource domains where such datasets were scarce (Specia et al., 2010).

The introduction of neural sequence-to-sequence (seq2seq) models significantly improved text simplification performance by leveraging distributed representations of words and contexts. Long Short-Term Memory (LSTM)-based architectures were among the first to achieve promising results, but they often suffered from exposure bias and difficulty handling long dependencies. This opened the door for more sophisticated architectures such as Transformers (Nisioi et al., 2017).

Transformer-based models like BART have become the state of the art for text simplification because of their bidirectional encoding and autoregressive decoding capabilities. They allow for rich contextual understanding while generating fluent, human-like text. Studies comparing BART to earlier seq2seq models demonstrate its superior ability to maintain semantic fidelity while reducing linguistic complexity, making it a strong foundation for real-world applications (Lewis et al., 2020).

Beyond simplification, researchers have emphasized the importance of supporting conceptual understanding, not just readability. Cognitive science research shows that analogy generation plays a critical role in helping learners transfer knowledge across domains, enabling them to connect new concepts with familiar experiences. This approach has been widely used in science education to explain abstract phenomena such as electricity, gravity, and quantum mechanics in terms that students can relate to (Richland & Simms, 2015).

Computational methods for analogy generation often use vector space models and semantic similarity measures to retrieve analogies from large datasets. While these methods can generate interesting comparisons, they are limited by the quality of their knowledge base and sometimes produce analogies that are too literal or lack pedagogical value, making them less helpful for human learning (Gentner & Forbus, 2011).

Recent advances combine retrieval mechanisms with generative models in what is known as Retrieval-Augmented Generation (RAG). This architecture retrieves relevant documents or examples from a knowledge base and uses them to guide the generative model's output, resulting in more factually accurate and contextually appropriate responses. This approach has been shown to improve coherence in tasks such as question answering, summarization, and educational content generation, suggesting its potential for analogy generation as well (Gao et al., 2023).

Despite these advancements, there is still a gap in tools that integrate text simplification and analogy generation into a single, user-friendly platform. Existing systems often focus on one aspect either simplifying text or generating analogies but do not combine them to support deeper comprehension. This creates an opportunity for developing a solution that not only simplifies language but also helps learners form conceptual links, thereby enhancing accessibility and knowledge transfer in educational and technical contexts.

MATERIALS AND METHODS

This section outlines the step-by-step approach used to develop the AI-driven platform for text simplification and analogy generation. It highlights the current systems, proposed system, and covers key processes such as data collection, model architecture, integration, and evaluation metrics. Each section is explained in detail to provide a clear understanding of the methodology and the reasoning behind each decision.

Analysis of the Current System

Modern text simplification and analogy generation systems leverage advanced NLP models to enhance comprehension of complex content. These systems primarily operate on the following mechanisms:

- i) **Rule-Based and Model-Driven Simplification:** Systems use predefined linguistic rules or models like BART and T5 for lexical and syntactic adjustments, converting dense text into readable forms while aiming to preserve meaning.
- ii) **Retrieval and Generative Analogy Creation:** Analogy systems source comparisons from datasets or generate new ones using embeddings and similarity metrics.
- iii) **Dataset-Dependent Training:** Models are trained on paired complex-simple texts or thematic data for paraphrasing and relational mapping.

Evaluation Through Metrics: Performance is assessed via automated scores like BLEU or human reviews, though often limited in capturing full semantic fidelity

Problems of the Current System

The main challenges with the existing text simplification and analogy generation systems include the following:

- i) **Semantic Integrity Loss:** Systems often lose essential information or alter meanings during simplification, particularly in technical texts.

- ii) **Restricted Adaptability:** Models trained on specific datasets struggle with niche domains, reducing accuracy and readability.
- iii) **Computational Requirements:** Advanced models like BART require significant resources, limiting real-time use.
- iv) **Static Training Data:** Reliance on fixed datasets hinders adaptation to evolving language or diverse content.
- v) **Limited Dataset Dependence for Analogies:** Systems produce narrow analogies due to domain-specific training.
- vi) **Lack of Context and Diversity:** Generated analogies are often general, failing to capture nuanced relationships.
- vii) **Data Collection Issues:** Preparing high-quality datasets is resource-intensive, impeding broad applicability.

Analysis of the Proposed System

The proposed text simplification and analogy generation platform integrates a pre-trained BART model with instruction-tuned large language models to deliver accurate, multi-level simplifications and context-aware analogies in real time. The text simplification module leverages BART to reduce linguistic complexity while preserving meaning, with an instruction-based LLM refining the output for readability and alignment with user-defined levels of simplicity. The analogy generation module transforms simplified text into vector embeddings, compares them with a thematic database using cosine similarity, and retrieves the most relevant sources, which are then combined with the original and simplified text by an LLM to generate clear, context-specific analogies. Both modules are unified within a single user-centric platform, allowing users to input complex text, select simplification levels, and choose analogy themes, with the final outputs presented in a clean, structured format that enhances comprehension and usability.

METHODOLOGY

The design and development of the proposed text simplification and analogy generation system follow a series of well-defined steps to ensure accurate, context-preserving outputs and a seamless user experience.

- i) **Data Collection and Preprocessing:** The first step involves gathering text-based data across multiple themes (e.g., education, sports) from public sources such as educational materials, encyclopedia entries, and open datasets. The data is cleaned, standardized, and converted into a format suitable for vector embedding generation before being stored in a vector database for efficient retrieval.
- ii) **Model Architecture and Training:** The system architecture is centered around two models: the text simplification model and the analogy generation model. Input text is first processed through a pre-trained BART model to produce a simplified version, which is then refined using an instruction-tuned LLM to ensure semantic consistency and alignment with the chosen simplification level.
- iii) **Analogy Generation Process:** The refined text is transformed into vector embeddings and compared against a pre-built vector database using cosine similarity to retrieve relevant thematic data. If the similarity threshold is met, the retrieved data is passed through the LLM to generate a contextually appropriate analogy tailored to the user's selected theme.
- iv) **Platform Development and Integration:** Both modules are integrated into a single platform designed for usability and responsiveness. The platform accepts user input, orchestrates the model pipelines, and displays well-formatted outputs including both the simplified text and any generated analogies in real time.

Evaluation and Optimization: System performance is measured using ROUGE-N for text overlap and BERTScore for semantic similarity. Iterative testing and parameter tuning are performed to ensure that both text simplification and analogy generation achieve high readability, relevance, and user satisfaction.

System Architecture

The proposed text simplification and analogy generation platform is designed with a modular, service-oriented architecture to ensure scalability, maintainability, and efficient processing. The system is composed of three key component:

i) *Text Simplification Module*: This module employs a two-stage pipeline. In the first stage, the pre-trained facebook/bart-large-cnn model performs summarization to restructure and simplify the input text while retaining its core meaning. In the second stage, the output is refined using Meta’s Llama 3.2B model through the Ollama framework, which provides graded simplification levels (1–3) to match user preferences.

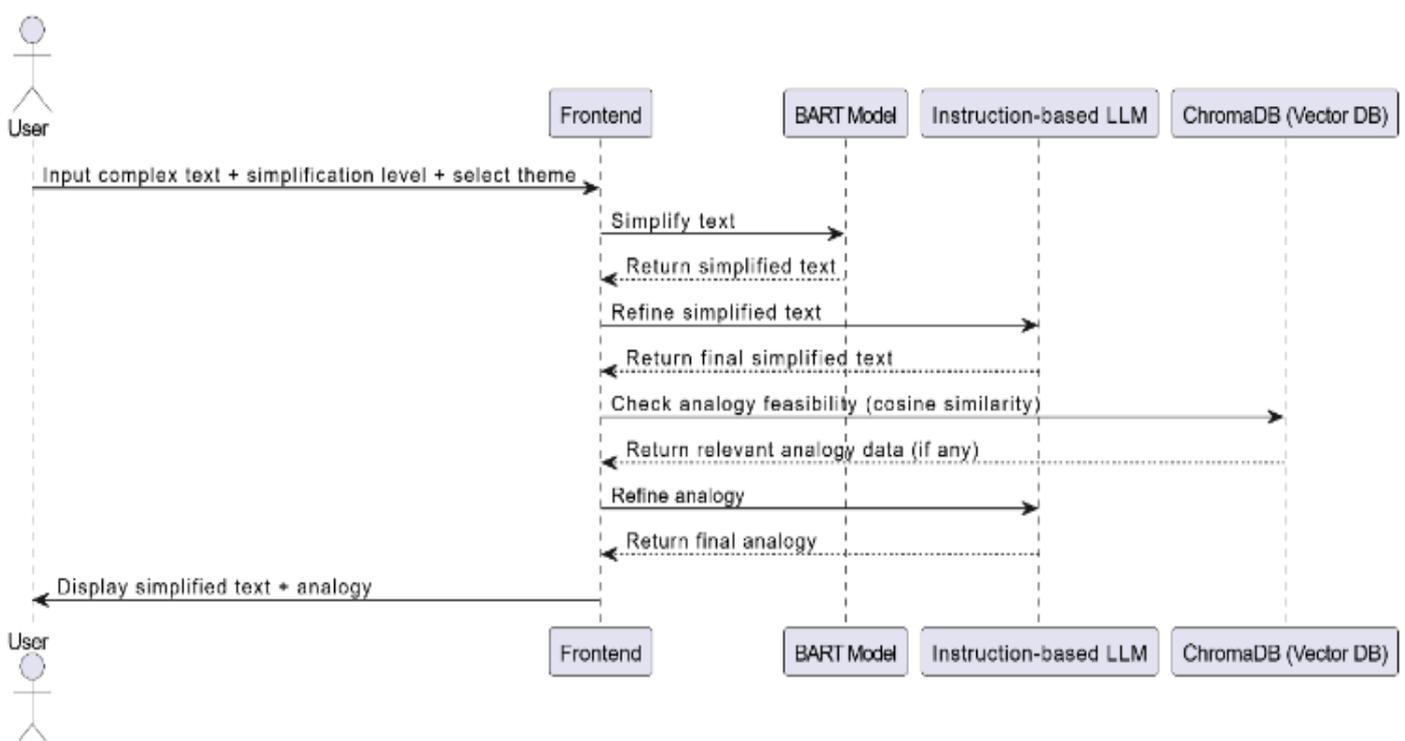
ii) *Analogy Generation Module*: This component is built on a four-stage Retrieval-Augmented Generation (RAG) pipeline. It begins with dataset preparation and embedding using mxbai-embed-large to create 512-dimensional semantic vectors, followed by theme-based retrieval from ChromaDB using cosine similarity. The retrieved data is then passed to Llama 3.2B for analogy synthesis, which leverages the BART-generated text for conceptual grounding. A feasibility filter ensures that only thematically relevant analogies are returned to the user.

iii) *Frontend–Backend Integration Layer*: A Flutter-based frontend provides an intuitive interface for user input and output display. This frontend communicates with a FastAPI backend over WebSocket connections, enabling low-latency, real-time processing for both text simplification and analogy generation requests.

System Design

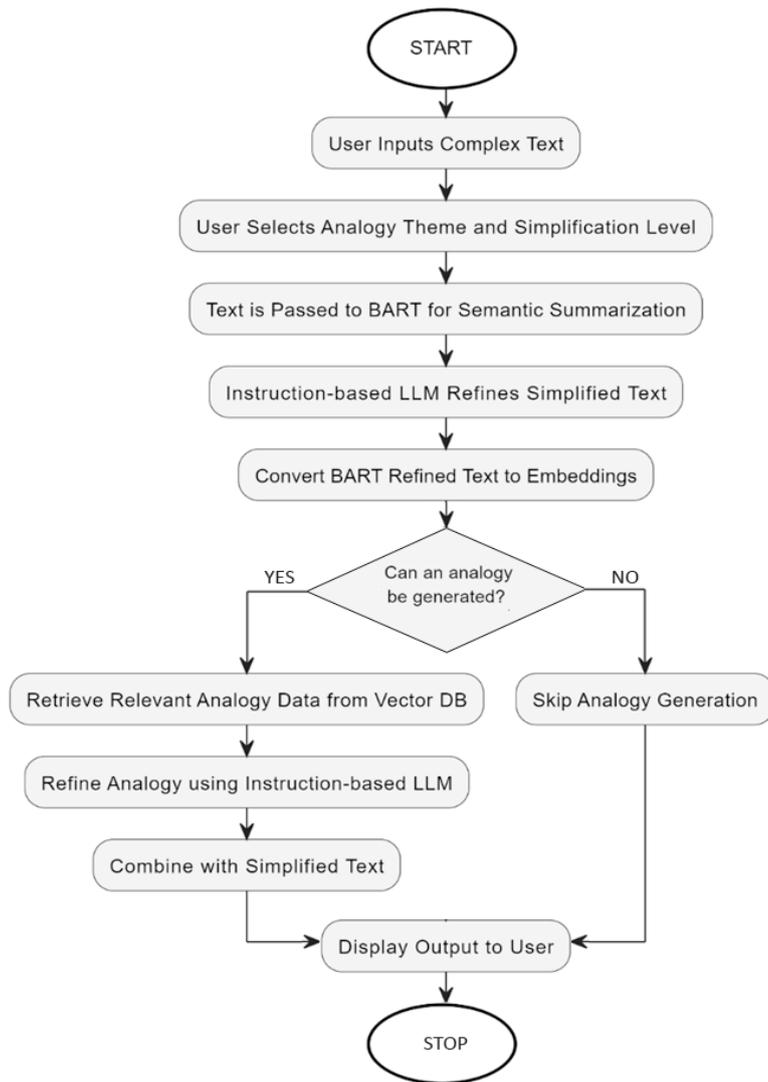
The system is built as a modular, user-friendly platform that combines text simplification and analogy generation in a single workflow. It uses a two-stage simplification pipeline with a pre-trained BART model for initial summarization and an instruction-tuned LLM for refinement, producing multiple levels of simplicity. Analogy generation is powered by a Retrieval-Augmented Generation (RAG) pipeline, which retrieves theme-relevant data from a vector database and synthesizes context-aware analogies. A Flutter frontend communicates with a FastAPI backend via WebSockets to enable smooth, real-time processing, ensuring that users receive clear, well-structured outputs quickly and efficiently

Fig. 1. Sequence Diagram



The sequence diagram shows the interaction flow within the text simplification and analogy generation platform. The process starts when the user submits a complex input text and selects an analogy theme through the frontend interface. The input is first simplified by the BART model, then refined by an instruction-tuned LLM to ensure clarity and semantic accuracy. The refined text is converted into vector embeddings and compared with thematic embeddings in ChromaDB using cosine similarity to check analogy feasibility. If the similarity threshold is met, the LLM generates or refines an analogy based on the retrieved data. Finally, both the simplified text and the generated analogy are sent back to the frontend and displayed to the user in a structured format.

Fig. 2. Activity Diagram



The activity diagram illustrates the step-by-step workflow of the platform, highlighting key decision points in text simplification and analogy generation. The process begins when the user inputs complex text and selects an analogy theme and simplification level. The input is first simplified by the BART model and then refined by an instruction-tuned LLM. The refined text is converted into vector embeddings and compared with thematic embeddings to determine if an analogy can be generated. If the similarity threshold is met, relevant analogy data is retrieved and refined; otherwise, the analogy step is skipped. Finally, the system displays the simplified text and any generated analogy to the user.

Proposed Programming Languages and Tools

Python was selected as the primary backend language due to its strong support for natural language processing and machine learning. Hugging Face’s Transformers library powers the BART model for text simplification, while ChromaDB serves as the vector database for storing and performing similarity searches on text

embeddings. MxBai is used to generate these embeddings, enabling fast and accurate analogy retrieval. For the frontend, Flutter was chosen to build a responsive, user-friendly interface that communicates with the backend via an API, allowing users to seamlessly access both text simplification and analogy generation features.

RESULTS, EVALUATION AND DISCUSSION

This section presents and evaluates the performance of the proposed text simplification and analogy generation system, focusing on both the quality of the simplified outputs and the relevance of the generated analogies. Evaluation was conducted using complex texts from diverse domains, including Wikipedia articles, with two representative test cases: Quantum Mechanics (Sports Dataset) and Cognitive Dissonance (Classroom Dataset). Simplification quality was assessed through ROUGE scores (ROUGE-1, ROUGE-2, ROUGE-L) for summarization performance and BERTScore (precision, recall, F1) to measure semantic similarity across three levels of simplification. Analogy generation was evaluated using BERTScore to ensure thematic relevance and conceptual alignment. Together, these results provide insight into the system's ability to preserve meaning, enhance readability, and generate contextually appropriate analogies

Results for Text Simplification

The text simplification pipeline processed the original texts through BART summarization followed by three levels of simplification. The results for each test are as follows:

1) Test 1: Sports Dataset (Quantum Mechanics)

Quantum mechanics was chosen as a test case because it involves complex scientific concepts about how matter and light behave at the atomic level. In this context, a successful simplification should preserve the core idea that quantum mechanics explains the behavior of tiny particles and energy.

i) *Original Text*: "Quantum mechanics is the fundamental physical theory that describes the behavior of matter and of light; its unusual characteristics typically occur at and below the scale of atoms. It is the foundation of all quantum physics, which includes quantum chemistry, quantum field theory, quantum technology, and quantum information science."

ii) *BART Summary*: "Quantum mechanics is the fundamental physical theory that describes the behavior of matter and of light. It is the foundation of quantum physics, which includes quantum chemistry, field theory, quantum technology, and quantum information science."

iii) Simplified Outputs:

Level 1: "Quantum mechanics is the fundamental physical theory describing how matter and light behave at an atomic scale. It's the base for all quantum physics fields, including chemistry, field theory, technology, and information science."

Level 2: "Quantum mechanics is an important theory that explains how matter and light behave. It's the basis for all quantum physics, which includes chemistry, field theory, technology, and information science"

Level 3: "Quantum mechanics is a way to understand how tiny things like atoms work. It explains how matter and light behave. This idea helps us with many areas like chemistry and technology."

1) Test 2: Classroom Dataset (Cognitive Dissonance)

Cognitive dissonance was chosen as a test case because it involves abstract psychological concepts about conflicting beliefs and how people resolve them. In this context, a successful simplification should preserve the core idea that people often hold opposing thoughts.

i) *Original Text*: "In the field of psychology, cognitive dissonance is described as a mental phenomenon in which people unknowingly hold fundamentally conflicting cognitions. Being confronted by situations that challenge this dissonance may ultimately result in some change in their cognitions or actions to cause greater alignment between them so as to reduce this dissonance."

ii) *BART Summary*: "Cognitive dissonance is a mental phenomenon in which people unknowingly hold fundamentally conflicting cognitions. Being confronted by situations that challenge this dissonance may ultimately result in some change in their cognitions or actions to cause greater alignment between them."

iii) *Simplified Outputs*:

Level 1: "Cognitive dissonance is a mental phenomenon in which people hold fundamentally conflicting cognitions without realizing it. Confronting situations that challenge this dissonance can lead to changes in their thoughts or actions, ultimately aligning them better"

Level 2: " Cognitive dissonance refers to when people hold conflicting ideas without realizing it. When confronted with situations that challenge these ideas, they may change their thoughts or actions to make them align better."

Level 3: " Cognitive dissonance is when people have conflicting thoughts. When they are faced with something that challenges these thoughts, it can make them change their mind or actions to fit better together."

Results for Analogy Generation

The analogy generation pipeline produced theme-specific analogies based on the simplified Level 1 outputs from each test. The results are as follows:

1) *Test 1: Sports Dataset (Quantum mechanics)*

Generated Analogy: "Quantum mechanics is like the rules in a football match: just as every football match follows a set of rules that shape an unpredictable game, quantum mechanics governs the tiny interactions of matter and light and how tiny particles behave in uncertain and unpredictable ways."

2) *Test 2: Classroom Dataset (Cognitive Dissonance)*

Generated Analogy: "Cognitive dissonance is like having two conflicting answers on a math test, when the teacher checks your work and finds that one answer contradicts another, it forces you to reconcile or reconsider the mistake and choose the correct solution."

Interface and Output Screenshots

Fig 3 displays the settings dialog, where users can select the desired simplification level and choose an analogy theme. Fig 4 shows the result after submitting text, highlighting both the simplified output and the generated analogy.

Fig. 3: Screenshot of Settings Dialog and Input

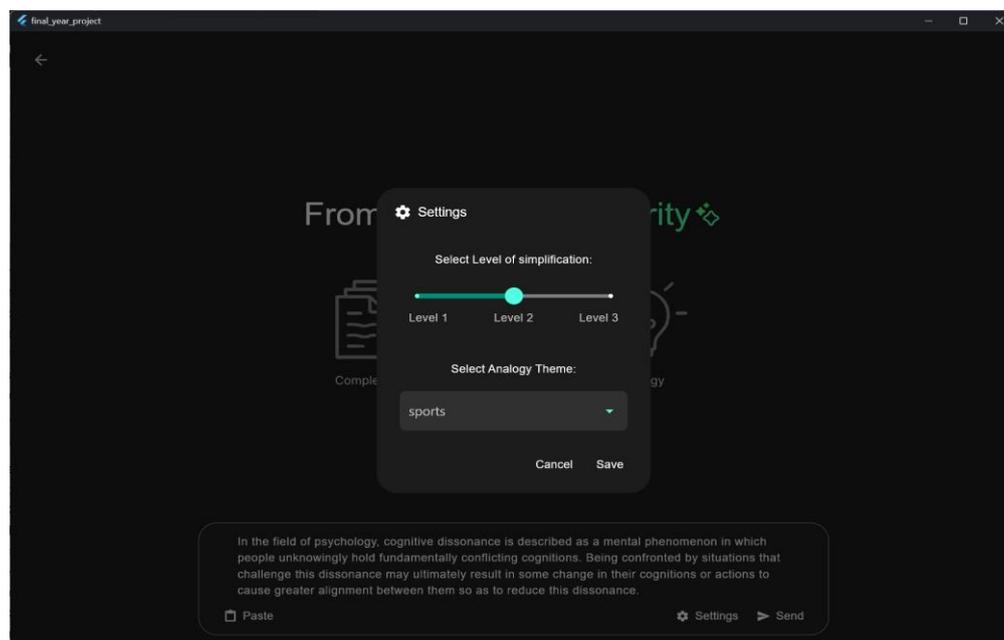
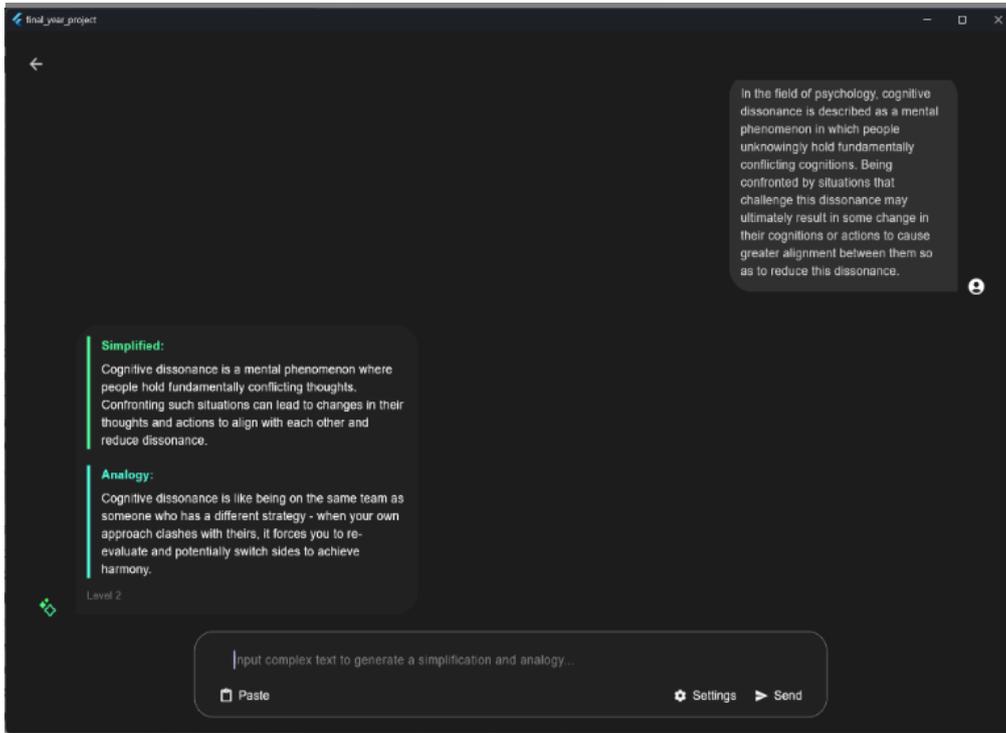


Fig. 4: Screenshot of Results



Evaluation

Fig 3 displays the settings dialog, where users can select the desired simplification level and choose an analogy theme. Fig 4 shows the result after submitting text, highlighting both the simplified output and the generated analogy.

Table I Evaluation Metrics For Test 1

Stage	Metrics
BART Summary	ROUGE-1: 0.8315, ROUGE-2: 0.8046, ROUGE-L: 0.8315
Simplification (Level 1)	F1: 0.9279 (Precision: 0.9822, Recall: 0.8794)
Simplification (Level 2)	F1: 0.8924 (Precision: 0.9485, Recall: 0.8426)
Simplification (Level 3)	F1: 0.8233 (Precision: 0.8825, Recall: 0.7716)
Analogy (Sports Theme)	F1: 0.7427 (Precision: 0.7452, Recall: 0.7403)

Table II Evaluation Metrics For Test 2

Stage	Metrics
BART Summary	ROUGE-1: 0.8387, ROUGE-2: 0.7912, ROUGE-L: 0.8387
Simplification (Level 1)	F1: 0.8742 (Precision: 0.9037, Recall: 0.8465)
Simplification (Level 2)	F1: 0.8638 (Precision: 0.9108, Recall: 0.8215)
Simplification (Level 3)	F1: 0.7984 (Precision: 0.8368, Recall: 0.7634)
Analogy (Classroom Theme)	F1: 0.7231 (Precision: 0.7382, Recall: 0.7087)

DISCUSSIONS

The evaluation results reveal that the proposed system performs well at preserving semantic integrity while improving readability, particularly at Level 1 simplification, which consistently achieved the highest F1 scores across both datasets. As simplification levels increased, a gradual decline in F1 scores was observed, reflecting the expected trade-off between simplicity and information retention, with Level 3 outputs being more accessible but less detailed. The BART summarization stage proved effective, as evidenced by high ROUGE scores, ensuring that key information was retained before further processing. Analogy generation also showed promising results, with BERTScore F1 scores above 0.7 for both datasets, producing meaningful and

contextually relevant analogies though some, like the sports analogy for quantum mechanics, were more abstract and required additional interpretation compared to the clearer classroom analogy. Overall, the system successfully balances clarity and semantic fidelity, highlighting the importance of selecting appropriate simplification depth and analogy framing based on the target audience.

CONCLUSION

This project successfully developed a functional platform for text simplification and analogy generation, leveraging a pre-trained BART model for summarization and simplification, alongside an instruction-tuned language model for theme-based analogy generation. Evaluation results demonstrated strong performance, with Level 1 simplification achieving the highest BERTScore F1 scores (0.9279 and 0.8742) on the Quantum Mechanics and Cognitive Dissonance datasets, indicating high semantic preservation and improved readability. While deeper levels of simplification led to a gradual decline in semantic precision, they still produced outputs accessible to broader audiences.

Analogy generation yielded F1 scores above 0.7, confirming the production of meaningful and contextually relevant analogies, though certain abstract domains required more interpretive effort. The integration of a Python backend with a Flutter frontend enabled real-time, user-friendly interaction.

Overall, the system provides a strong foundation for educational and accessible NLP tools. Future work may focus on improving analogy relevance through fine-tuned models and expanding dataset diversity to better support domain-specific comprehension across various user groups.

RECOMMENDATIONS

Although the project successfully achieved its objectives, several challenges were encountered, including limited access to large, real-world datasets and computational constraints during model development. While these were partially mitigated, the following areas are recommended for future work:

- i) **Dataset Expansion:** Extend the training and evaluation datasets to cover more domains and contexts. A broader dataset would enable the platform to produce richer, more relevant analogies and handle a wider range of text complexities.
- ii) **Algorithm Refinement:** Enhance the simplification pipeline by introducing feasibility checks and reinforcement learning techniques to prevent over-simplification while preserving core meaning.
- iii) **Improved Analogy Evaluation:** Incorporate multi-dimensional similarity metrics beyond cosine similarity and fine-tune thresholds for domain-specific embeddings to ensure that generated analogies are more contextually aligned and intuitive.
- iv) **Cloud-Based Deployment:** Migrate the processing pipeline to cloud-hosted or online LLMs to improve computational efficiency, scalability, and support real-time responses for larger datasets.
- v) **User Training and Documentation:** Develop comprehensive documentation and training resources for educators, students, and developers to encourage platform adoption and gather feedback for continuous improvement.

REFERENCES

1. Seidenberg, M. S. (2013). The science of reading and its educational implications. *Psychological Science in the Public Interest*, 14(1), 1–54. <https://doi.org/10.1177/1529100612453577>
2. Crossley, S. A., Allen, D. B., & McNamara, D. S. (2011). Text readability and intuitive simplification: A comparison of readability formulas. *Reading in a Foreign Language*, 23(1), 1–20. <https://doi.org/10.64152/10125/66657>

3. Richland, L. E., & Simms, N. (2015). Analogical reasoning and the transfer of learning. *Educational Psychology Review*, 27(4), 599–614. <https://doi.org/10.1007/s10648-015-9321-4>
4. Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 7871–7880. <https://doi.org/10.18653/v1/2020.acl-main.703>
5. Lazaridou, A., Peysakhovich, A., & Lewis, M. (2021). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Proceedings of the 37th International Conference on Machine Learning*, 129–139. <https://arxiv.org/abs/2005.11401>
6. Carr, N. (2010). *The Shallows: What the Internet Is Doing to Our Brains*. W.W. Norton & Company.
7. Richland, L. E., Holyoak, K. J., & Stigler, J. W. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist*, 47(3), 189–203. <https://doi.org/10.1080/00461520.2012.667065>
8. Martínez, P., Ramos, A., & Moreno, L. (2024). Exploring large language models to generate easy to read content. *Frontiers in Computer Science*, 6, 1394705. <https://doi.org/10.3389/fcomp.2024.1394705>
9. Siddharthan, A. (2014). A survey of research on text simplification. *International Journal of Applied Linguistics*, 165(2), 259–298. <https://doi.org/10.1075/itl.165.2.06sid>
10. Specia, L., Gasperin, C., & Santos, D. (2010). Translating from complex to simplified sentences. In *Proceedings of Coling 2010* (pp. 1–9). Springer. https://doi.org/10.1007/978-3-642-12320-7_5
11. Nisioi, S., Štajner, S., Ponzetto, S. P., & Dinu, L. P. (2017). Exploring neural text simplification models. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Short Papers)*, 85–91. <https://doi.org/10.18653/v1/P17-2014>
12. Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(2), 177–192. <https://doi.org/10.1002/wcs.1336>
13. Gentner, D., & Forbus, K. D. (2011). Computational models of analogy. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(3), 266–276. <https://doi.org/10.1002/wcs.105>
14. Gao, Y., et al. (2023). Retrieval-augmented generation for large language models: A survey. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2312.10997>