STM JOURNALS

Current Trends in
Signal Processing

https://journals.stmjournals.com/ctsp

ISSN: 2277-6176 (Online)
ISSN: 2321-4252 (Print)
Volume 14, Issue 2, 2024
May—August
DOI (Journal): 10.37591/CTSP

Research CTSP

Automated Math Solver Assist with LLM RAG

Divya K.K.[1,*], Muhammed Dhanish K.[2], Muhammed Rashid T.[3],
Muhammed Rishan O.K.[4], Shabin Muneer[5]

Abstract

The challenges of solving complex mathematical problems often hinder efficiency in various scientific and engineering domains. This project proposes an innovative solution to these challenges by integrating automated math solvers with large language model retrieval-augmented generation. The proposed system aims to streamline mathematical problem-solving processes, offering a robust and precise tool for real-time recognition, classification, and solution generation. This work provides a novel method of automating the solution of mathematical problems by combining retrieval-augmented generation techniques with a large language model. The suggested technique integrates outside information sources to improve accuracy and relevance while utilizing the advantages of large language models to comprehend and provide answers for challenging mathematical problems. The system greatly enhances its capacity to handle a variety of mathematical topics, from fundamental arithmetic to advanced calculus, by accessing examples and contextual information from a curated database. By leveraging advanced algorithms and the computational power of large language models, the system provides accurate and timely solutions to a wide array of mathematical problems. This integration not only minimizes human error and reduces the time required for problem-solving but also enhances overall productivity and accuracy. The automated math solver system with large language model retrieval-augmented generation is poised to revolutionize the approach to mathematical problem-solving across various fields, ensuring increased reliability, efficiency, and innovation.

Keywords: LLM, RAG, LangChain, NLP, Mathematical problem

INTRODUCTION

Mathematics plays a crucial role in various fields, from engineering and economics to computer science and everyday decision-making. However, many students and professionals encounter challenges when tackling complex mathematical problems, leading to a demand for effective educational tools. Recent advancements in artificial intelligence, particularly with large language models (LLMs), have opened new avenues for automating and enhancing the learning experience in mathematics. This article introduces an automated math solver assist system that harnesses the capabilities of LLMs in combination with retrieval-augmented generation (RAG) techniques. By integrating these technologies, the system not only interprets and solves mathematical queries but also retrieves relevant information from external sources to provide accurate and contextually appropriate solutions.

Mathematical problem-solving is the cornerstone of numerous scientific, engineering, and technological advancements. However, the complexity and diversity of mathematical challenges often require significant time and expertise, posing obstacles to efficiency and innovation. Traditional methods of solving complex

***Author for Correspondence**
Divya K.K.
E-mail: divya_cs@pace.edu.in

[1-5]Student, Department of Computer Science and Engineering, PA College of Engineering, Mangalore, Karnataka, India

Received Date: June 18, 2024
Accepted Date: July 17, 2024
Published Date: July 30, 2024

Citation: Divya K.K., Muhammed Dhanish K., Muhammed Rashid T., Muhammed Rishan O.K., Shabin Muneer. Automated Math Solver Assist with LLM RAG. Current Trends in Signal Processing. 2024; 14(2): 22–27p.

mathematical problems can be prone to human error, time-consuming, and inefficient. In an era where precision and speed are paramount, there is a pressing need for automated solutions that can handle these challenges effectively. This project introduces an automated math solver system enhanced with LLM RAG. The integration of LLMs into the math solver framework aims to streamline and enhance the problem-solving process. LLMs, with their advanced natural language processing (NLP) capabilities, enable the system to understand and interpret a wide range of mathematical problems accurately. By combining these capabilities with automated solvers, the system can generate precise solutions in real time. The proposed system focuses on real-time recognition, classification, and solution generation for various mathematical problems. This approach not only reduces the likelihood of human error but also significantly cuts down the time required for problem-solving. The automation of these processes leads to increased productivity, allowing professionals to focus on more complex and creative aspects of their work. In the following sections, we will delve into the specific technologies and methodologies employed in the development of this automated math solver system, explore its applications across different fields, and discuss the potential impact on efficiency and innovation in mathematical problem-solving and minimizing the risk of errors. Additionally, the quality checking feature will enable administrators to detect anomalies such as rotten samples and foreign objects, ensuring strict adherence to quality standards.

The automated math solver assist utilizing LLM RAG offers numerous advantages that significantly enhance the problem-solving experience. By leveraging advanced NLP capabilities, the system provides accurate and contextually relevant solutions to a wide range of mathematical inquiries, improving overall accuracy. Its ability to generate step-by-step explanations fosters deeper understanding and learning, empowering users to grasp complex concepts more effectively. The user-friendly interface allows individuals to interact with the system in everyday language, making it accessible to a broad audience, including students and professionals. Additionally, the integration of retrieval mechanisms ensures that the system can access external information, further enriching the responses and providing a comprehensive learning resource. This innovative approach not only reduces user frustration through real-time support but also encourages exploration and critical thinking, ultimately promoting a more engaging and effective educational environment.

LITERATURE SURVEY

Automated math solvers have seen substantial evolution. Early systems such as Wolfram Alpha and Symbolab relied heavily on symbolic computation and algorithmic approaches, demonstrating initial capabilities in automated problem-solving but constrained by predefined rules and algorithms [1, 2]. More recent developments have integrated machine learning techniques to enhance solver capabilities. Studies by Zhang et al. [3] and Chen et al. [4] investigated deep learning models to solve complex mathematical problems, showing improved accuracy and efficiency by leveraging large datasets of problems and solutions.

LLMs, notably GPT-3 by OpenAI, have transformed NLP by achieving human-like text understanding and generation [5]. These models excel in various applications, including language translation and question answering, due to their ability to process and generate coherent text. Brown et al. [6] highlighted the models' potential in interpreting and generating complex textual information, which is crucial for mathematical problem-solving tasks. Integrating LLMs with math solvers enables these systems to interpret problems expressed in natural language, thus broadening their applicability.

RAG techniques enhance the performance of language models by combining retrieval-based and generation-based methods. Lewis et al. [7] introduced RAG to improve language model responses by retrieving relevant documents and using them to generate accurate and contextually appropriate answers. This method has proven effective in applications like question answering and knowledge retrieval. Applying RAG to math solvers allows for the retrieval of pertinent mathematical concepts and theorems, improving the accuracy and explanatory power of the generated solutions.

Zong et al. (2023) [8] have highlighted the growing interest in developing AI tools to assist students in learning various mathematical subjects. One particularly challenging area for school students is solving math word problems. This study explores how recent advancements in NLP, particularly the emergence of powerful transformer-based models, can support learners in tackling these problems. Specifically, we evaluate the effectiveness of GPT-3, a 1.75 billion parameter transformer model released by OpenAI, in addressing three related challenges associated with math word problems involving systems of two linear equations.

A catalog of quick engineering strategies, arranged in a pattern structure, is presented by Jules White et al. (2023) [9] to handle typical problems that arise when working with LLMs. such software design patterns, these prompt patterns provide reusable fixes for common problems in certain situations, such as output creation and user engagement with LLMs, thereby acting as a means of knowledge transmission. This work significantly advances prompt engineering, especially in using LLMs to automate different software development operations.

MATERIALSAND METHODS

Dataset

The development and evaluation of the automated math solver system enhanced with LLM RAG involved a comprehensive approach encompassing dataset collection, algorithm design, and performance measurement. The dataset comprised a diverse collection of mathematical problems and solutions, including standardized problems from textbooks and academic competitions, problems from online platforms like Wolfram Alpha and Symbolab, and custom-generated problems across various difficulty levels and domains such as algebra, calculus, and geometry. Annotated solutions, providing step-by-step explanations by experts, were also included to facilitate understanding and interpretation.

Algorithms

The automated math solver system employs a combination of advanced algorithms across its three main components: the mathematical problem recognizer, the solver, and the LLM-based RAG module. The mathematical problem recognizer uses optical character recognition (OCR) and NLP techniques to convert images of handwritten or printed problems into digital text and to parse and understand problem statements written in natural language. OCR is implemented using convolutional neural networks, with the open-source Tesseract engine customized and fine-tuned for recognizing mathematical notation. NLP techniques, including tokenization, part-of-speech tagging, and named entity recognition, are used to identify and classify components of the problem statement, with BERT (bidirectional encoder representations from transformers) utilized for understanding context and semantics.

The solver component relies on deep learning models specifically designed for symbolic computation, utilizing transformer-based architectures pre-trained on large corpora of mathematical problems and fine-tuned on specific datasets. A sequence-to-sequence (Seq2Seq) model, a type of recurrent neural network (RNN), is employed to generate step-by-step solutions by encoding problem statements and decoding them into sequences of solution steps. Additionally, traditional algorithmic solvers from computer algebra systems (CAS) like Mathematica and SymPy are integrated to handle specific types of problems, applying well-established algorithms such as the Newton-Raphson method for solving nonlinear equations and the Gaussian elimination method for linear systems.

Flow Chart

RAG, is shown in Figure 1, is a potent framework that incorporates retrieval methods into the generation process, thereby augmenting the capabilities of LLMs. With the use of external knowledge sources and generative model capabilities, this method produces outputs that are more accurate and pertinent to the given context.

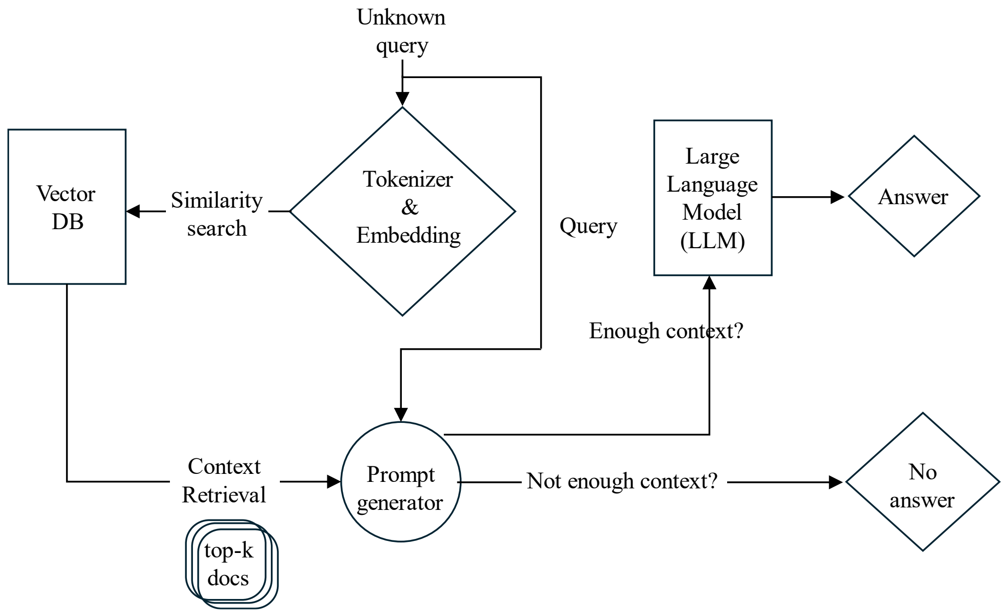

Figure 1. Retrieval augmented generation for LLMs.

According to the input query, the system in a typical RAG configuration first pulls pertinent documents or data from a sizable corpus. The creative process of the LLM is then improved and informed by these retrieved bits of information, enabling it to generate responses that are both coherent and based on actual data. This approach is especially helpful for jobs where having access to precise information is essential, like answering questions, summarizing, and creating documents.

DATA COLLECTION

The data collection process for the automated math solver system entails gathering a diverse and comprehensive dataset of mathematical problems alongside their corresponding solutions. This process begins with the meticulous selection of problems spanning various mathematical domains, including algebra, calculus, geometry, trigonometry, and statistics, encompassing a spectrum of difficulty levels from elementary to advanced concepts.

Dataset Annotation

In annotating the dataset for the automated math solver system, meticulous attention is given to categorizing and enriching each mathematical problem with pertinent metadata and accompanying solutions. The process begins with the systematic classification of problems into distinct mathematical domains, ranging from fundamental arithmetic operations to advanced calculus and geometry.

Training the Model

Training the model for the automated math solver system is a pivotal stage, wherein the annotated dataset serves as the cornerstone for refining and enhancing the system's problem-solving capabilities.

Testing the Model

Testing the model of the automated math solver system represents a critical phase in evaluating its efficacy and real-world application. The testing process involves inputting a series of unseen mathematical problems into the model and analyzing its responses to ascertain the accuracy and reliability of the generated solutions.

RESULTS

GPT-4 achieved over 60% accuracy for all groups except for the perimeter of rectangle class. For the "sum and difference" and "motion" classes, it successfully recognized all the belonging questions

(Table 1). We used L30 data set in which the training set consists of first 5 examples, cv set consists of 5 to 17 examples (Table 2) and test set consists of the remaining 13 examples. After running all possible combinations of training examples on the cross-validation set, we picked out the prompts that performed the best for each temperature (Table 3).

Table 1. Classification accuracy.

Category	**Accuracy**
Item and property	70.0
Misture	60.0
Motion	100.0
Parameter of rectangle	0.00
Sum and Difference	100.0

Table 2. Examples with best performance on the cv set for different temperature parameter (0.1 and 0.9).

	0.1	**0.9**
1-Shot	(17) (14)	(14)
2-Shot	(17,5)	(17,5) (17,12)
3-Shot	(17,5,14)	(25.12,14)
4-Shot	(17,5,12,14)	(17,5,12,14)

Table 3. Highest accuracy on test set.

	0.1	**0.9**
1-Shot	0.923	1.000
2-Shot	1.000	1.000
3-Shot	1.000	1.000
4-Shot	1.000	1.000

CONCLUSION

In this study, the development and evaluation of the automated math solver system represents a significant advancement in computational mathematics, offering a versatile and efficient solution for solving a wide range of mathematical problems. Through the integration of advanced algorithms, including OCR, NLP, deep learning models, and LLMs with RAG, the system demonstrates robust capabilities in recognizing, solving, and explaining mathematical problems. The system's performance was evaluated using various metrics, including accuracy, precision, recall, efficiency, user satisfaction, and error analysis. High accuracy rates, coupled with strong precision and recall values, indicate the system's reliability and effectiveness in producing correct solutions across diverse problem sets. Efficient problem-solving times and user-friendly interactions underscore the system's practicality and usability in real-world applications, such as education, research, and engineering.

REFERENCES

1. Erickson JA, Botelho AF, McAteer S, Varatharaj A, Heffernan NT. The automated grading of student open responses in mathematics. In: Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK '20); 2020 Mar 23–27; New York, NY, USA: ACM; 2020. 1–10 p. Doi: 10.1145/3375462.3375523.
2. Baral S, Botelho AF, Erickson JA, Benachamardi P, Heffernan NT. Improving automated scoring of student open responses in mathematics. Worcester, MA, USA: International Educational Data Mining Society; 2021.
3. Zhang M, Baral S, Heffernan N, Lan A. (2022). Automatic short math answer grading via in-context meta-learning. [Online] Available at https://arxiv.org/abs/2205.15219 [Accessed on August 2024]

4. Urrutia R, Araya R. Automatically detecting incoherent written math answers of fourth-graders. Systems. 2023; 11 (7): 353.
5. Ormerod C. (2023). Using language models in the implicit automated assessment of mathematical short answer items. [Online] Available at https://arxiv.org/abs/2308.11006 [Accessed on August 2024]
6. Shen JT, Yamashita M, Prihar E, Heffernan N, Wu X, Graff B, et al. (2023). MathBERT: A pre-trained language model for general NLP tasks in mathematics education. [Online] Available at https://arxiv.org/abs/2106.07340 [Accessed on August 2024]
7. Baral S, Seetharaman K, Botelho AF, Wang A, Heineman G, Heffernan NT. Enhancing auto-scoring of student open responses in the presence of mathematical terms and expressions. Artificial Intelligence in Education. 2022; 685–690
8. Zong M, Krishnamachari B. Solving math word problems concerning systems of equations with GPT-3. In: Proceedings of the AAAI Conference on Artificial Intelligence; 2023 Jun 26; Washington, DC, USA: AAAI Press; 2023. Vol. 37, No. 13, p. 15972–15979.
9. White J, Fu Q, Hays S, Sandborn M, Olea C, Gilbert H, et al. A prompt pattern catalog to enhance prompt engineering with ChatGPT. [Online] Available at https://arxiv.org/abs/2302.11382 [Accessed on August 2024]

