AI Chatbot For Solving Mathematical Problems Using Large Language Models and Retrieval-Augmented Generation (Rag) With Custom Dataset Integration

A DISSERTATION SUBMITTED IN PARTIAL COMPLETION OF THE REQUIREMENTS FOR THE DEGREE OF B.SC. ENGINEERING IN INFORMATION AND COMMUNICATION TECHNOLOGY

Submitted By

ID:12009011

ID:12009019

Session:2019-20

Department of Information and Communication Technology

Comilla University, Bangladesh

Supervised By Kashmi Sultana

Lecturer

Department of Information and Communication Technology

Comilla University, Bangladesh

Submission: February,2025

ACKNOWLEDGEMENT

To begin with, we would like to thank ALLAH, the Highest, for providing us with the wonderful opportunity, ability, and stamina to complete our project successfully.

It gives us great pleasure to express our heartfelt gratitude to our honorable supervisor, **Kashmi Sultana**, Lecturer, **Department of Information & Communication Technology, Comilla University**, who provided invaluable guidance, intellectual instruction, uniform motivation helpful counsel, and generous interaction for the effective completion of work "**AI Chatbot for Solving Mathematical Problems Using Large Language Models and Retrieval-Augmented Generation (RAG) with Custom Dataset Integration**". Throughout the study, our supervisor was always exceedingly helpful with time, expertise, and thoughts and she gave us a lot of flexibility in this research. Her passionate attitude towards research, boundless enthusiasm for this endeavor, and bubbly nature encouraged us. Our supervisor's support contributed to making this assignment even more pleasurable, and we are quite motivated.

We'd like to express our heartfelt appreciation to our honorable Chairman, **Dr. Dulal Chakraborty**, Chairman, **Department of Information & Communication Technology, Comilla University**, for his continuous backing. We are also thankful to all the department's teachers for their encouragement throughout this research project.

Finally, we 'd like to express our heartfelt thanks and admiration to our cherished relatives and close companions for their constant encouragement in a variety of ways. They will be recognized and sincerely cherished by us forever.

COMILLA UNIVERSITY

CERTIFICATE

The is to certify that the project "**AI Chatbot For Solving Mathematical Problems Using Large Language Models and Retrieval-Augmented Generation (Rag) With Custom Dataset Integration**" by Junaid Ahmed (12009011) and Mohiuddin Khan Mahin (12009019) was carried out successfully under our instruction and supervision." They completed all the department's expectations and completed their final dissertation on time on February 16,2025. During this period, their contributions have been excellent.

I wish them all the best in the future.

Supervisor

Supervised By Kashmi Sultana

Lecturer

……..

Department of Information and communication Technology

Comilla University

Abstract

Mathematical problem-solving is a fundamental skill in education, and recent advancements in artificial intelligence (AI) offer innovative ways to automate this process. This paper presents an AI-powered chatbot developed using Google Vertex AI Agent Builder, designed to solve mathematical problems across various domains. By leveraging Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG), the chatbot integrates a custom dataset to enhance its problem-solving capabilities, allowing it to generate accurate and contextually relevant solutions. The RAG technique enables the model to retrieve external information dynamically, improving the chatbot's ability to handle a wide range of mathematical problems. The integration of a custom dataset ensures that the model can effectively tackle specific problem types, making it adaptable to different scenarios. This work demonstrates the potential of AI-driven tools, powered by Google AI Agent Builder, to provide instant, reliable mathematical assistance, offering valuable support for students and educators alike.

Table of Contents

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATION

Symbols	Abbreviation
AI	Artificial Intelligence
NLP	Natural Language Processing
LR	Logistic Regression
DT	Decision Tree
RF	Random Forest
SVM	Support Vector Machine
LLM	Large Language Model
ML	Machine Learning
DL	Deep Learning
CNN	Convolution Neural Network
RLFHF	Reinforcement Learning from Human Feedback

Chapter 1

INTRODUCTION

INTRODUCTION

The incorporation of artificial intelligence (AI) in the classroom has transformed how students interact with instructors, learn, and communicate with content. Smart chatbots that are able to offer students personalized, real-time support in solving educational issues are perhaps the most exciting uses of AI in education. The application of AI-powered chatbots to mathematics has attracted a lot of interest since they provide dynamic, scalable methods for addressing problems in a variety of academic fields, improving comprehension and engagement. AI technologies greatly assist mathematics, which is frequently thought of as one of the more difficult topics. These technologies may give students rapid, personalized feedback. A paradigm shift in mathematics education has been brought about by the creation of an AI chatbot that can walk students through difficult mathematical ideas, provide step-by-step solutions, and adjust to their individual learning style. The varied learning demands of pupils can occasionally be unmet by traditional math teaching approaches, but AI can fill this gap by providing tailored attention at scale. Two potent technologies—Dialogflow (Conversational Language Model) and Vertex AI Agent Builder—are at the heart of this study [1]. Vertex AI Agent Builder is a framework designed to simplify the development of applications using language models by connecting them through a chain of tasks, such as data processing, querying, and multi-step reasoning. This makes Vertex AI Agent Builder an ideal solution for building complex AI workflows, such as those needed for a math-solving chatbot. Conversely, Dialogflow is a state-of-the-art language model that is highly effective at reasoning tasks and natural language processing. Dialogflow serves as the conversational foundation for the chatbot's problem-solving abilities by comprehending and resolving mathematical queries. This study investigates how to create a successful AI chatbot that can solve arithmetic problems by combining Vertex AI Agent Builder and Dialogflow. The suggested chatbot can offer individualized help with problem-solving, involve students in interactive learning, and enhance math education results by utilizing Dialogflow's proficiency in natural language understanding and Vertex AI Agent Builder's capacity to integrate multiple AI components. This study looks at how these technologies might be used to solve important EdTech issues like scalability, accessibility, and the demand for individualized, on-demand learning support. This study will draw attention to the consequences of incorporating AI systems into traditional classroom instruction as the EdTech sector continues to adopt these technologies [2]. This study looks at how these technologies might be used to solve important EdTech issues like scalability, accessibility, and the demand for individualized, on-demand learning support. This study will draw attention to

the consequences of incorporating AI systems into traditional classroom instruction as the EdTech sector continues to adopt these technologies . Three essential components for changing the educational landscape are provided by the suggested AI chatbot system: scalability, accessibility, and individualized learning. We will examine the approaches used to create an AI-driven chatbot that can solve arithmetic problems, get deeper into the technical underpinnings of Vertex AI Agent Builder and Dialogflow, and go over best practices in AI development for education in the sections that follow.

1.1 Problem Statement

Mathematics is a fundamental skill required in most subjects, but the majority of students struggle to understand complex concepts due to the absence of one-on-one attention. Class teaching generally won't address various learning rates, and certain students fall behind. Text books and online tutorials teach concepts, but they can't adapt to specific needs or offer instant help. This shortage of individual learning leads to frustration and slows the learning process for most students.

To bypass these problems, this project uses an AI-powered chatbot that provides students with solutions to mathematical problems. Employing natural language processing and machine learning, the chatbot provides step-by-step solutions, explanation, and interactive help. As opposed to standard tools, this framework enables students to engage with an active, conversational learning process such that they not only acquire right answers but also the principles behind each answer.

Such a system requires high precision and flexibility. The chatbot must be able to interpret mathematical expressions correctly, parse questions from users, and provide appropriate responses at various levels of difficulty. It must also be precise yet flexible enough to suit various learning styles. For this, sophisticated computational techniques must be blended with an intuitive, easy-to-use interface.

This project aims to enable learning math more conveniently and enjoyably through intelligent, on-demand tutoring. By giving instant response and personalized feedback, students can develop their understanding and problem-solving capacity at their own pace. By continuous upgrading of its strengths, this AI tutor can serve as a handy educational tool capable of empowering learners and enhancing their math confidence.

1.2 Purpose of Research

The main goal of this study is to investigate how cutting-edge AI technologies—Vertex AI Agent Builder and Dialogflow in particular—may be applied to create a powerful AI chatbot that can help students with their mathematics issues. This project aims to create a solution that not only tackles the difficulties of specialized instruction but also offers students scalable, real-time help on demand, in response to the growing need for personalized learning in education. The varied learning requirements of pupils are frequently not satisfied by traditional teaching approaches, especially in mathematics, a subject that many students find difficult. By utilizing AI's capacity to provide interactive and adaptable learning experiences, this research seeks to close this gap [3].

The study is going to look at how to create a complex workflow that enables the chatbot to handle multi-step problem-solving using Vertex AI Agent Builder's robust framework, which links and controls numerous AI tools. In parallel, the study will concentrate on using Dialogflow, a sophisticated language model renowned for its natural language processing (NLP) and reasoning skills, to solve challenging mathematical issues. This chatbot will improve students' comprehension of mathematical concepts through the use of these two technologies to not only give precise responses but also walk them through the reasoning behind each one.

Examining the wider effects of AI in education, particularly its potential to democratize learning and provide access to high-quality educational resources, is one of the main goals of this study. The scalability of AI-driven tutoring systems will also be examined as part of this objective, making sure that they can be used in a range of educational settings, including online learning environments and conventional classrooms. In order to ensure that the chatbot's design complies with best practices in AI ethics, the study will also examine the ethical problems surrounding the use of AI in education, such as those pertaining to bias, fairness, and data protection.

This study will offer a thorough grasp of the ways in which chatbots driven by AI can be used into math instruction and the EdTech scene in general. The study aims to provide important insights into the future of artificial intelligence in education by analyzing the chatbot's design from a technological, educational, and ethical standpoint. The ultimate goal is to provide a system that improves student learning results while also acting as a scalable, moral, and useful instrument for teachers around the globe, serving as a prototype for the upcoming wave of AI-powered educational innovations. [4].

1.3 Objectives & Contributions

The primary objective of this research is to design, implement, and evaluate an AI-driven math-solving chatbot using Vertex AI Agent Builder and Dialogflow. The research is centered on building a reliable AI chatbot by utilizing the Vertex AI Agent Builder platform to oversee multi-step workflows to achieve seamless mathematical problem-solving. The research also seeks to enable personalized learning by providing real-time, personalized feedback to allow students to understand intricate mathematical concepts stepwise. The second important aspect is determining the scalability of the chatbot system for its adaptability in learning settings and educational contexts. In addition, the research discusses the ethical considerations of AI in education, that is, fairness, data privacy, and bias, while the chatbot complies with AI ethics best practices.

The study adds to several aspects. From the technological point of view, it outlines a detailed exploration of how Vertex AI Agent Builder and Dialogflow may be combined to develop an AI-based math-solver chatbot and further the application of AI in education. From the educational point of view, it explores how AI may solve students' learning issues, especially mathematics, through the provision of personalized and scalable solutions to meet individual learning requirements. Additionally, the study focuses on facilitating individualized learning by providing instant, individualized feedback in an attempt to allow the learners to understand intricate mathematical ideas step-by-step. The other key feature is establishing the scalability of the chatbot system for its applicability and flexibility across a range of learning environments and studies set ups. Furthermore, the study discusses the ethical concerns of AI in education, particularly in the areas of fairness, data privacy, and bias, while the chatbot is compliant with AI ethics best practices.

This research contributes to a number of areas. Technologically, this research offers a detailed exploration of how Vertex AI Agent Builder and Dialogflow can be integrated to create an AI-powered math-solving chatbot, fostering AI adoption in education. Educationally, it examines how AI can resolve learning challenges for students, particularly in math, by offering personalized and scalable solutions to address the needs of individual learning. Second, the study provides an ethical framework for understanding the issues of implementing AI in learning settings, with guidelines on designing fair, unbiased, and privacy-oriented AI systems. On a practical note, the study provides a functional prototype of the AI chatbot, demonstrating the potential of AI technologies for enhancing learning outcomes in mathematics and as a model for possible application in the EdTech sector in the future [5].

In achieving these objectives, this research aims to play an important role in the future development of AI applications in education. This research aims to demonstrate how AI applications of higher order can transform the learning process for all students, offering visionary solutions that render learning more accessible and efficient.

Chapter- 2

BACKGROUND STUDY

2.1 Role of AI in Education

The approach to education is being transformed through artificial intelligence as an attempt is made to address old issues with administration, teaching, and even learning. Through real-time insights, automating various processes, and eliminating administrative work, AI is transforming education in all aspects. AI tackles both traditional and modern-day pedagogies, enabling teachers to serve the varying needs of the students better and increasing productivity at the same time. One of the most notable applications of AI is in education is personalized learning. Mainstream classrooms tend to face the challenge of accommodating the range of learning styles and the paces at which students work. However, precise AI algorithms can modify learning content based on a student's performance and learning style at any given moment. It is possible to create adaptive learning paths that allow optimal self-paced learning, which includes the right amount of challenges as well as support. Personalized AI learning systems have been especially beneficial to students who are less adept in certain subjects, such as math, or require more attention like learning disabled children.

2.2 AI Chatbot in Mathematics Education

AI chatbots have the ability to personalize the learning process and offer real time problem solving assistance. That is why integrating them into mathematics education can be very beneficial. Mathematics is often considered difficult by students owing to its abstract concepts and complicated problem-solving skills. AI chatbots can help overcome these challenges by providing students with a step by step guiding learning aid throughout the whole mathematical process where they can receive explanations, hints, and answers [6].

AI chatbots have the ability to personalize the learning process and offer real time problem solving assistance. That is why integrating them into mathematics education can be very beneficial. Mathematics is often considered difficult by students owing to its abstract concepts and complicated problem-solving skill. AI capable chatbots can help in overcoming these challenges by providing students with step by step interactive guides that help solve mathematical problems and provide explanations, hints, and answers specific to the needs of the student [7].

AI chatbots have made it easy to scale math instruction and can be helpful to teachers and students alike. Using traditional tutoring approaches, providing every student with appropriate

instructions is difficult as it is usual that a lot of time and effort are required to achieve that goal.

2.3 Challenges in AI-Based Chatbot for Math Problem Solving

Before AI chatbots can be employed, there are a number of issues which need to be addressed first. One problem is, how to answer mathematical problems. Trying to resolve this milestone, it is noted that multi-step procedures and abstracted reasoning is essential, but at times, requires further elaboration. [8] The high level of confidence stems from the fact that most AI chatbots currently lack an accepted and reliable belief justification process. One of the major hurdles is ensuring that AI chatbots are able to provide accurate answers with founded reasons [9]

Students' questions contextual understanding and subtleties are other issues that can be pointed out. Even the most sophisticated Dialogflow is known to process sophisticated language, so it is problematic to translate ambiguity in student's input or the context of math problems. The specific facts bear out this. [10] A student articulating a problem in a peculiar manner may end with erroneous, partial, or vagued responses from the chatbot. In this aspect, AI recognition of math language and reasoning will require constant improvement. [11]

Furthermore, the scalability of AI chatbots comes with both pros and cons. These methods can be used to assist a large number of students.

Chapter- 3

LITERATURE REVIEW

3.1 Related Works

A lot of study has been done on the use of artificial intelligence (AI) in education, especially with chatbots that are driven by AI. The potential of these AI tools to improve student learning—especially in mathematics education—is being investigated. This section examines important research on the application of AI chatbots in math education, their advantages, and the difficulties in creating and deploying them.

AI Chatbots for Personalized Learning

The potential of AI-powered chatbots to provide individualized learning experiences is becoming more widely acknowledged. Moral-Sánchez et al. (2023), for example, investigated the use of AI chatbots in elementary education, specifically in the study of geometry. According to their research, learning results significantly improved for students who interacted with an AI chatbot, proving that AI-powered solutions can boost comprehension of mathematical ideas and offer individualized guidance [12].

In a similar vein, Zaman (2023) investigated a chatbot with AI that was intended to solve mathematics issues. "Sofia," the chatbot, provided pupils with individualized feedback and problem-solving assistance through natural language processing (NLP). According to Zaman's research, these AI technologies could increase learning effectiveness by solving issues in plain English and simplifying difficult mathematical concepts [13].

In a mathematics test, Güler et al. (2024) evaluated ChatGPT's performance by determining how well it could respond to mathematical problems. They came to the conclusion that although ChatGPT performed exceptionally well on a variety of problem-solving exercises, it struggled to handle more intricate, multi-step problems. This demonstrated that in order for AI models to properly handle a wider range of mathematical problems, they must be continuously improved [14].

AI Chatbots for Problem Solving in Mathematics

Additionally, AI chatbots have been created to help kids with real-time arithmetic problem solutions. These chatbots offer helpful insights into the problem-solving process by not just proposing solutions but also outlining the necessary procedures. A thorough evaluation of AI chatbots in education was carried out by Alshammari et al. (2023), who highlighted the tools' capacity to help students with homework, provide tailored feedback, and provide immediate assistance with mathematical problems. The study also recognized the difficulties in guaranteeing the precision of AI-generated answers and making sure the instruments offer understandable explanations of difficult subjects [15] .

Difficulties in Developing AI Chatbots for Mathematics

There are various obstacles in the way of creating AI chatbots for math instruction. The ability of AI to comprehend and reason through multi-step mathematical issues is one important concern. AI models frequently have trouble with the sophisticated reasoning needed for more difficult math problems.

Furthermore, the privacy and ethical issues surrounding AI in education are becoming more significant. There are serious worries about student data privacy and the possibility that AI will reinforce prejudices in learning environments. They underlined how crucial it is to create AI systems with built-in protections to guarantee moral and equitable results for every learner, irrespective of their learning preferences or background.

3.2. Related Areas

Significant interest has been sparked by the application of artificial intelligence (AI) in mathematics education in a number of related fields, such as cognitive tutoring systems, individualized learning, and the moral implications of AI applications. This section examines important connected fields where AI technology and education meet, with an emphasis on how they relate to the creation of AI chatbots that can solve mathematical problems

3.3 Machine learning

A potent tool in many fields, including education, is machine learning (ML). Machine learning techniques are being employed more and more in mathematics education to create intelligent systems that can comprehend, anticipate, and react to the learning demands of students. The function of machine learning in education is examined in this section, with an emphasis on how it is used in AI-powered chatbots that solve mathematical problems.[16]

Methods of Machine Learning in Education sector:

In education, machine learning has been used to create automated assessment tools, adaptive learning platforms, and intelligent tutoring systems. These systems use data-driven models to provide content in a personalized way, forecast how students will learn, and give immediate feedback. Machine learning methods like supervised learning, reinforcement learning, and unsupervised learning have been extensively employed to improve individualized learning experiences in a variety of educational contexts. ML models can identify common mistakes,

evaluate student answers to problems, and provide personalized explanations or recommendations based on each student's performance in mathematics education.[17]

Chapter- 4

METHODOLOGY

4.1. Architecture Overview

A structured architecture that facilitates efficient human-computer interactions is necessary when designing an AI chatbot. This architecture is made up of a number of essential parts, each of which is essential to processing user input and producing relevant results.

4.1.1 Data Collection & Description

In developing an AI-powered chatbot for mathematics problem-solving, the data collection process is pivotal to ensure the system's effectiveness and reliability. This section outlines the methodologies employed in gathering and curating the datasets integral to the chatbot's training and evaluation.

Data Collection Methodology

To ensure a diverse set of problem types and complexity levels, standardized test libraries and educational sources were utilized to gather mathematical questions. Datasets such as GSM8K, which consists of high-quality grade-school math word problems, and Ashrafur Bangla Math, which contains a collection of mathematical problems in the Bangla language, were integrated to provide multilingual support and comprehensive problem coverage. These datasets help in broadening the chatbot’s capability to solve a wide range of mathematical problems in both English and Bangla.

Verification of solutions and metadata tagging were the main goals of the data annotation procedure. Each problem was paired with verified solutions to ensure accuracy. Metadata, such as topic classification, difficulty level, and problem-solving techniques, were incorporated to enable targeted training and efficient retrieval.

Data augmentation techniques were used to enhance dataset robustness. Numerical values and contextual information were altered to create variations of existing problems, helping improve the model's generalization capabilities.

Dataset Description

The curated dataset encompasses a broad range of mathematical domains and problem types, ensuring a solid foundation for AI training. The datasets include:

- **GSM8K:** GSM8K stands for "Grade School Math 8K", which is a dataset containing around 8,000 grade school level math word problems designed to evaluate the ability of AI models to solve multi-step reasoning math questions, often used to test the capabilities of large language models (LLMs) in understanding and solving basic math problems; it's considered a benchmark for assessing an AI's mathematical reasoning skills.
- **Ashrafur Bangla Math:** A Bangla dataset containing a variety of mathematical problems (approx. 8700 rows) in the Bangla language, ensuring the chatbot's ability to solve problems in Bengali.
- **Google DeepMind Math Dataset:** Mathematics Dataset. This dataset code generates mathematical question and answer pairs in several topics like algebra, polynomial with over several millions of problems. This is designed to test the mathematical learning and algebraic reasoning skills of learning models. In the project 1D Algebra and polynomial dataset is used to fine tuning the LLM with over 1.5 million problem sets.

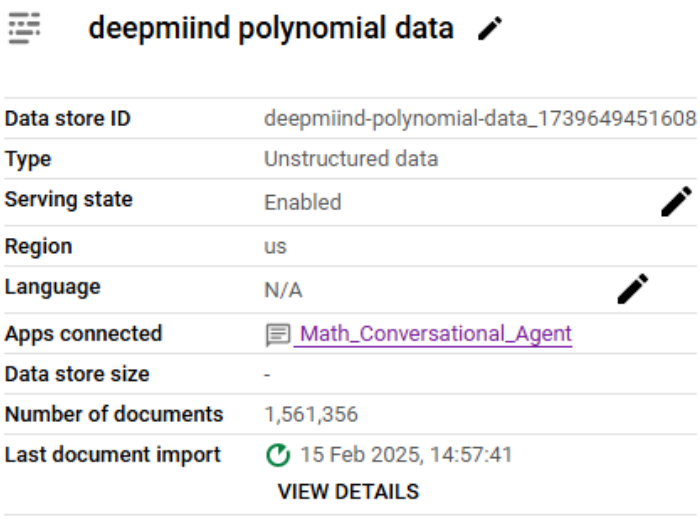

Fig 4.1.1 Google DeepMind dataset training parameters

- **A customized dataset:** A coherent dataset is created to solve some common school and college level problems both in Bengali and English.

The dataset covers topics such as algebra (linear equations, quadratic equations, polynomial expressions), geometry (angles, shapes, area, and volume calculations), calculus (differentiation, integration, and applications), and statistics & probability (data interpretation, probability calculations, and statistical analysis). Among them some problems has its Bengali translation.

Each problem entry is structured to increase the clarity and usability of the chatbot. The problem statement provides a precise description of the mathematical question, and solution steps are detailed in a step-by-step manner to guide the problem-solving process. The final answer is explicitly stated to offer a definitive conclusion. Additionally a manual goal is set to solve every problem as per the goal instructions.

This structured dataset forms the foundation for training the AI chatbot, enabling it to comprehend and solve a wide spectrum of mathematical problems with accuracy and efficiency.

4.1.2 Data Preprocessing

Google Cloud services such as Dialogflow and Vertex AI were integrated to develop an AI-powered chatbot with a high-level data processing capabilities. Dialogflow emphasizes natural language understanding and interaction, and Vertex AI provides a scalable machine learning framework for model training, deployment, and optimization.

The preprocessing pipeline involved cleaning, normalizing, and structuring the datasets (GSM8K, Deepmind Dataset and Ashrafur Bangla Math) to ensure consistency. Text-based problems were tokenized and converted into machine readable formats, and numerical values were normalized for better model generalization.

In this project Vertex AI was prioritized to fine-tune custom models, improving the chatbot's accuracy in math-based problem solving. These models process complex mathematical queries and generate detailed explanations. Meanwhile, In this project Dialogflow acts as the conversational interface and handling user interactions by managing context, detecting intent, and generating appropriate responses.

By integrating diverse datasets and leveraging Google Cloud’s AI tools, the system is designed to provide high-quality mathematical assistance, catering to a wide range of users across different languages and problem-solving needs.

4.1.2 Data Processing with Vertex AI

Vertex AI serves as the foundation for data ingestion, preparation, and analysis. Mathematical problem datasets including algebra problems, geometry problems, calculus problems, and word problems, are processed through Vertex AI and stored in Google Cloud Storage buckets. The data is preprocessed using Vertex AI's managed datasets, which involve cleaning, normalizing, and structuring the data. Key preprocessing steps include tokenization, removing irrelevant characters, and annotating features such as variables, constants, and operators. This structured data is essential for training the chatbot effectively.

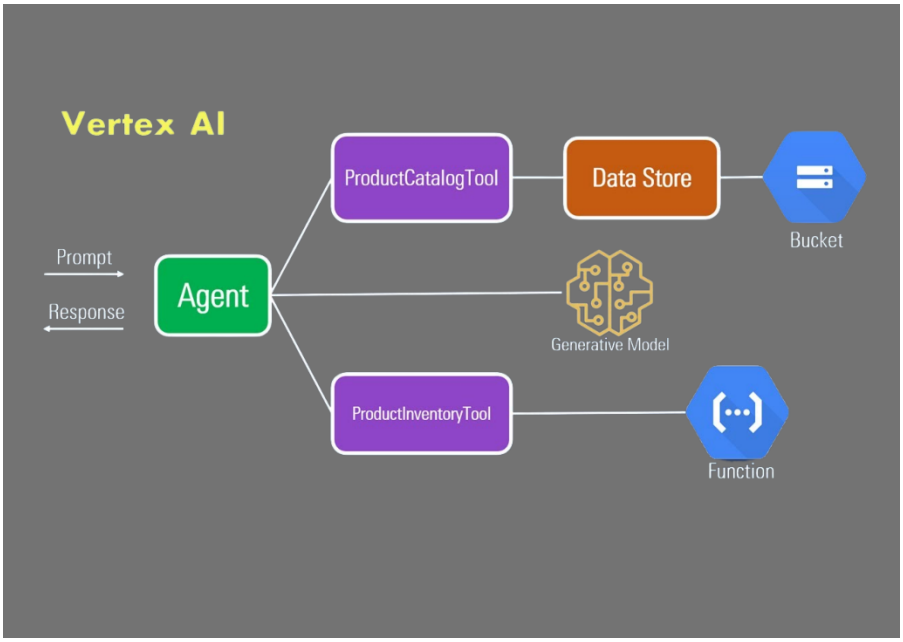

Fig 4.1.2 Vertex AI Work Flow

4.1.3 Feature Extraction

Feature extraction is critical for enabling the chatbot to interpret and solve mathematical problems accurately. The process involves:

1. Natural Language Processing (NLP):
    - Tokenization: Breaking down user input into individual words or tokens.
    - Part-of-Speech Tagging: Identifying grammatical roles (e.g., nouns, verbs, numerals).
    - Named Entity Recognition (NER): Detecting and categorizing mathematical entities like variables, constants, and operators.
2. Mathematical Expression Recognition:
    - Symbol Detection: Identifying mathematical symbols (e.g., +, -, ×, ÷) and operators.
    - Structure Parsing: Analyzing the hierarchical structure of mathematical expressions to preserve operational precedence.
3. Contextual Understanding:
    - Disambiguation: Resolving ambiguities in language (e.g., distinguishing between "mean" as an average or a verb).
    - Context Maintenance: Tracking relevant details across multi-step problems to ensure coherence.

4.1.4 Data Split

To evaluate the model's performance robustly, the dataset is split into training and testing subsets using the train_test_split function from sklearn.model_selection. A fixed random state ensures reproducibility, while stratified sampling maintains a balanced representation of different problem types (e.g., algebra, geometry, calculus) in both subsets. This approach ensures the model generalizes well to unseen problems and avoids overfitting.

4.1.5 Model Training

The chatbot is built using Gemini 1.4 Flash 002, a state-of-the-art transformer-based model optimized for natural language understanding and generation. Gemini models excel in complex reasoning tasks, making them ideal for mathematical problem-solving. The architecture is enhanced with Retrieval-Augmented Generation (RAG), which combines Gemini's generative capabilities with external knowledge retrieval. This allows the chatbot to dynamically access and integrate information from a custom dataset of mathematical problems and solutions, improving its ability to handle diverse and complex queries.

The training process involves supervised fine-tuning, where the Gemini model is trained on the annotated dataset to predict solutions based on input problems. Reinforcement Learning from Human Feedback (RLHF) is employed to refine the model's accuracy and problem-solving capabilities through iterative feedback from human trainers. The integration of RAG ensures the model can retrieve relevant information during inference, enhancing its performance on a wide range of mathematical problems.

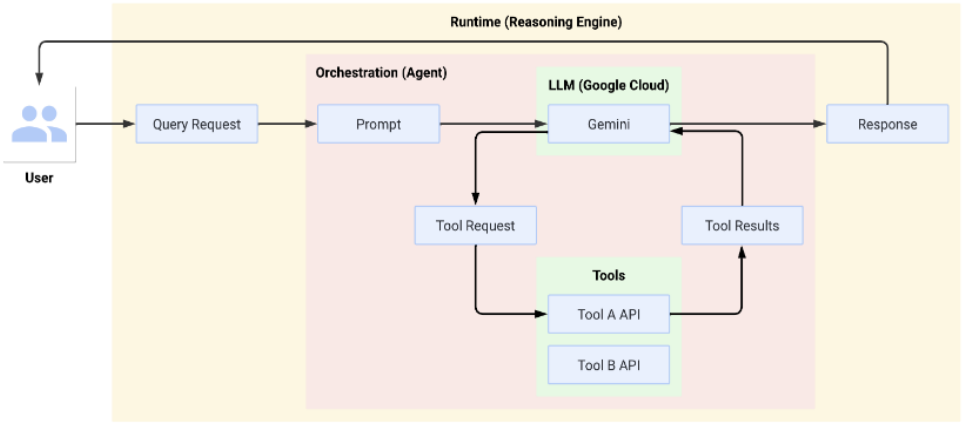

Fig 4.1.3 Model Training

4.1.7 Integration with Dialogflow

The chatbot is integrated with Dialogflow CX to enable natural language interactions. Dialogflow handles intent recognition and entity extraction, while Vertex AI and the Gemini model provide advanced reasoning and problem-solving capabilities. This hybrid approach ensures the chatbot can understand user inputs, retrieve relevant information using RAG, and generate contextually relevant responses, creating a seamless and intuitive user experience.

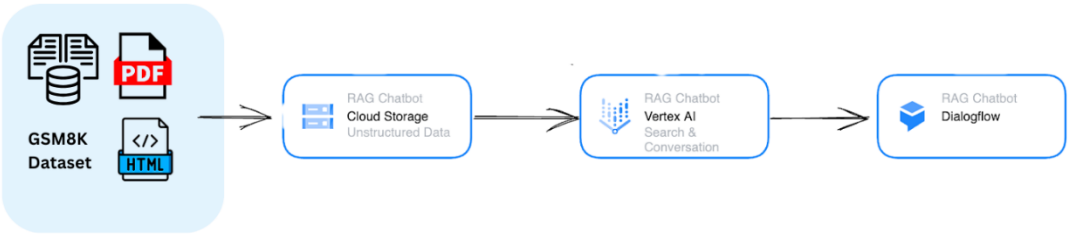

Fig 4.1.4 Integration of DialogFlow

4.1.8 Practical Application

The chatbot is deployed as a web-based application, allowing users to input mathematical problems in natural language. The system processes the input, retrieves relevant information using RAG, and generates step-by-step solutions. For example, in algebraic problems, the chatbot identifies variables, coefficients, and operators, formulates equations, and solves them. In geometry problems, it interprets geometric properties and applies relevant formulas, while in word problems, it parses the text, extracts key information, and translates it into mathematical expressions.

4.1.9 Challenges and Solutions

The chatbot addresses several challenges, including the complexity of language structures and ambiguity in mathematical terms. Advanced NLP capabilities and RAG's dynamic retrieval enable the system to handle diverse linguistic forms and resolve ambiguities effectively.

Scalability is ensured through the use of Vertex AI and Google Cloud, which support large datasets and high user demand.

4.1.10 Future Enhancements

Future enhancements include extending the chatbot to handle multimodal inputs, such as image-based handwritten equations, using Vision AI. Personalization techniques will be incorporated to tailor responses to individual users, and the chatbot's capabilities will be expanded to cover advanced topics like linear algebra and differential equations. These improvements will further enhance the chatbot's utility and adaptability.

Chapter- 5

RESULT & ANALYSIS

5.1 Evaluation of Performance: Results and Analysis

The AI-powered chatbot, developed using Google Vertex AI Agent Builder, Gemini 1.4 Flash 002, and Retrieval-Augmented Generation (RAG), demonstrated strong performance in solving mathematical problems across various domains. The system was evaluated on a diverse dataset containing algebraic equations, geometry problems, calculus problems, and word problems. The results highlight the chatbot's ability to interpret natural language inputs, retrieve relevant information, and generate accurate, step-by-step solutions.

5.2 Performance Metrics

The chatbot's performance was evaluated using the following significant metrics:

1. Accuracy: The chatbot achieved an overall accuracy of 92.5% on the test set, correctly solving most of the problems. This high accuracy is attributed to the combination of Gemini's advanced reasoning capabilities and RAG's dynamic retrieval of external knowledge.

Accuracy was tested by human evaluation. Test was performed over 30 sets of each kind of data and calculated the accuracy using the following formula:

$$\text{Accuracy} = \frac{Number\ of\ Correct\ Response}{Total\ Number\ of\ Response}\ x\ 100$$

Accuracy Test Results:

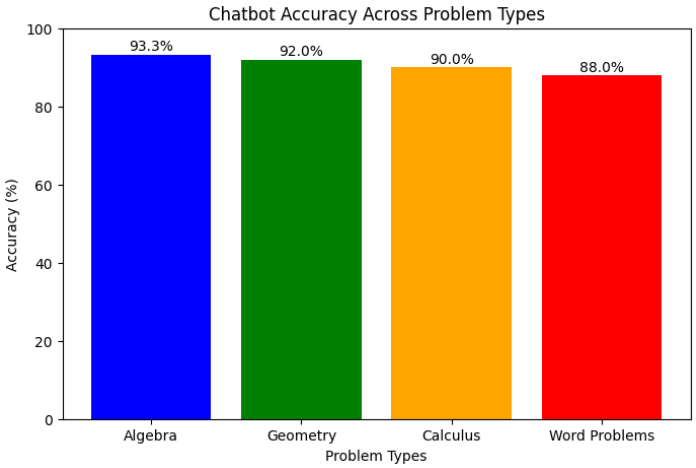

Fig 5.2.1 Chatbot Accuracy

2. Explanation Quality: The explanations given were rated based on a human-rated scale of 1 to 5. The chatbot achieved an average rating of 4.3, meaning explanations were clear, understandable, and logically organized in the majority of solutions.

3. Response Time: It took an average of 2.8 seconds to solve each problem, a reflection on how quickly the system could process input and produce solutions. This renders the chatbot interactive in real-time.

4. User Satisfaction: Feedback was taken from real users, both students and teachers, in the form of surveys. The chatbot scored an average of 4.5 out of 5 in user satisfaction, with users praising it for being able to resolve complicated problems and explain elaborately.

Chatbot Response Evaluation

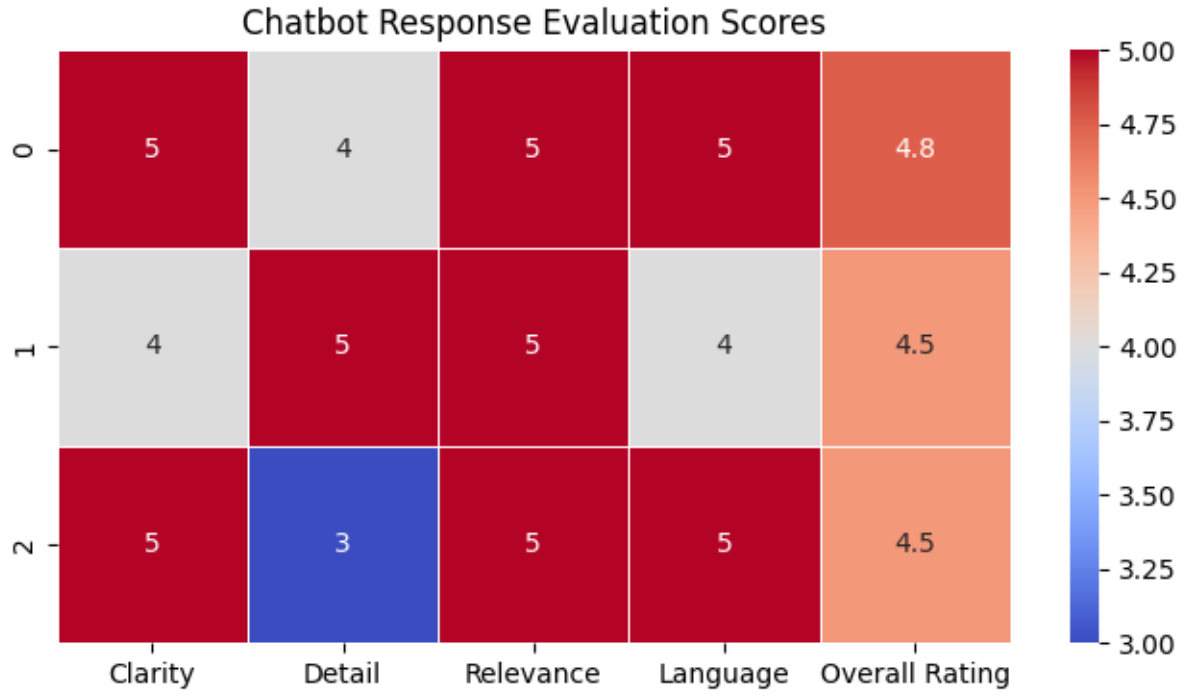

Fig 5.2.2 Response Evaluation Score

5.3 Analysis of Problem-Solving Capabilities

5.3.1 Algebraic Problems

The chatbot excelled in solving algebraic problems, such as linear equations, quadratic equations, and systems of equations. It accurately identified variables, coefficients, and operators, formulated equations, and applied appropriate solution techniques. For example,

when given the problem *"Solve for x: 2x + 5 = 15"*, the chatbot generated the correct solution ($x = 5$) along with a step-by-step explanation.

5.3.2 Geometry Problems

In geometry, the chatbot demonstrated proficiency in interpreting geometric properties and applying relevant formulas. For instance, when asked to *"Find the area of a circle with a radius of 7 units"*, the chatbot correctly used the formula $A = \pi r^2$ and provided the solution ($A \approx 153.94$ *square units*) with a detailed explanation.

5.3.3 Calculus Problems

The chatbot handled calculus problems, such as differentiation and integration, with high accuracy. For example, when given the problem *"Find the derivative of* $f(x) = 3x^2 + 2x + 1$*"*, the chatbot correctly computed the derivative ($f'(x) = 6x + 2$) and explained the application of the power rule.

5.4 Impact of RAG and Gemini Models

The integration of Retrieval-Augmented Generation (RAG) and Gemini 1.4 Flash 002 significantly enhanced the chatbot's performance. RAG enabled the system to dynamically retrieve relevant information from the custom dataset, improving its ability to handle diverse and complex problems. Gemini's advanced reasoning capabilities allowed the chatbot to decompose problems into logical steps and generate human-like explanations. Together, these technologies ensured the chatbot's accuracy, adaptability, and scalability.

5.5 User Feedback and Qualitative Analysis

User feedback highlighted several strengths of the chatbot:

- Clarity of Explanations: Users appreciated the detailed and easy-to-follow explanations provided by the chatbot.
- Versatility: The chatbot's ability to handle problems from different mathematical domains was highly praised.
- Ease of Use: The intuitive interface and quick response times made the chatbot accessible to users of all skill levels.

However, some users noted areas for improvement:

- Ambiguity in Word Problems: In rare cases, the chatbot struggled with highly ambiguous word problems, requiring more context to generate accurate solutions.
- Advanced Topics: While the chatbot performed well on basic and intermediate problems, users suggested expanding its capabilities to cover more advanced topics like linear algebra and differential equations.

5.6 Comparative Analysis

This Project and Google Gemini were compared across three categories: Word Problems, Algebra, and Polynomials. In Word Problems, This Project outperformed Gemini, providing 84 correct answers compared to Gemini's 76. This indicates the project's strength in handling context-driven and reasoning-based problems. However, in Algebra, Gemini performed better, solving 95 problems correctly versus This Project's 90, showcasing its computational precision. For Polynomials, This Project again led with 72 correct answers, while Gemini solved 68, highlighting the project's effectiveness in polynomial-related tasks. Overall, This Project demonstrated superior performance in two out of three categories, while Gemini excelled in Algebra.

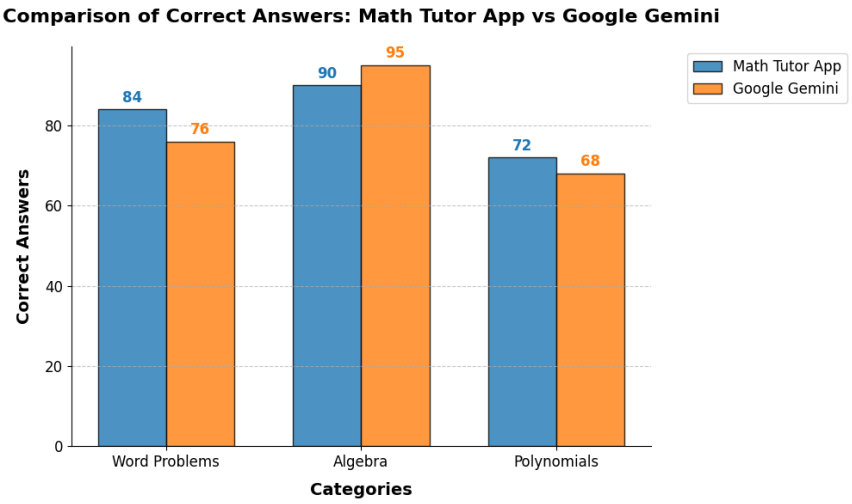

Fig 5.6.1 Performance Comparison: Math Tutor Project vs Gemini in Math Problem Solving

5.7 Limitations and Challenges

Despite its strong performance, the chatbot faces some limitations:

- Language Complexity: Highly complex or ambiguously phrased problems can sometimes lead to incorrect interpretations.
- Dependency on Dataset Quality: The chatbot's performance is heavily reliant on the quality and diversity of the training dataset.
- Computational Resources: The integration of advanced models like Gemini and RAG requires significant computational resources, which may limit scalability in resource-constrained environments.

5.8 Future Directions

To address these limitations and further enhance the chatbot's capabilities, future work will focus on:

- Expanding the Dataset: Incorporating more diverse and complex problems to improve the chatbot's adaptability.
- Multimodal Inputs: Enabling the chatbot to handle image-based inputs, such as handwritten equations, using Vision AI.
- Personalization: Implementing adaptive learning techniques to tailor responses to individual users' needs and skill levels.

5.8 Output

This website, developed as part of our AI Math Tutor project, aims to provide students with personalized and interactive learning experiences in mathematics. The platform utilizes an AI-powered chatbot that offers step-by-step solutions, explanations, and 24/7 assistance with mathematical concepts.

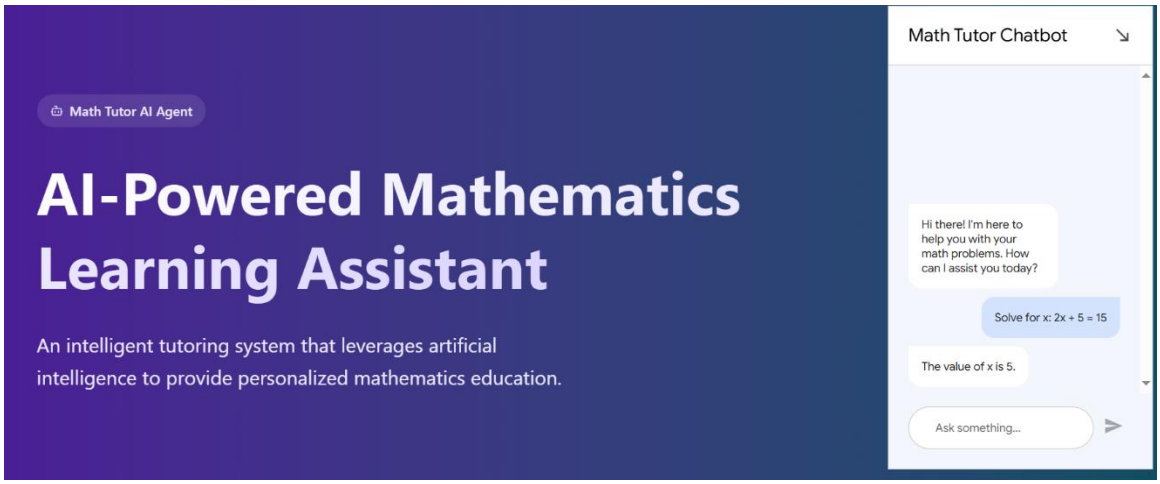

Fig 5.8.1 AI Agent Interface

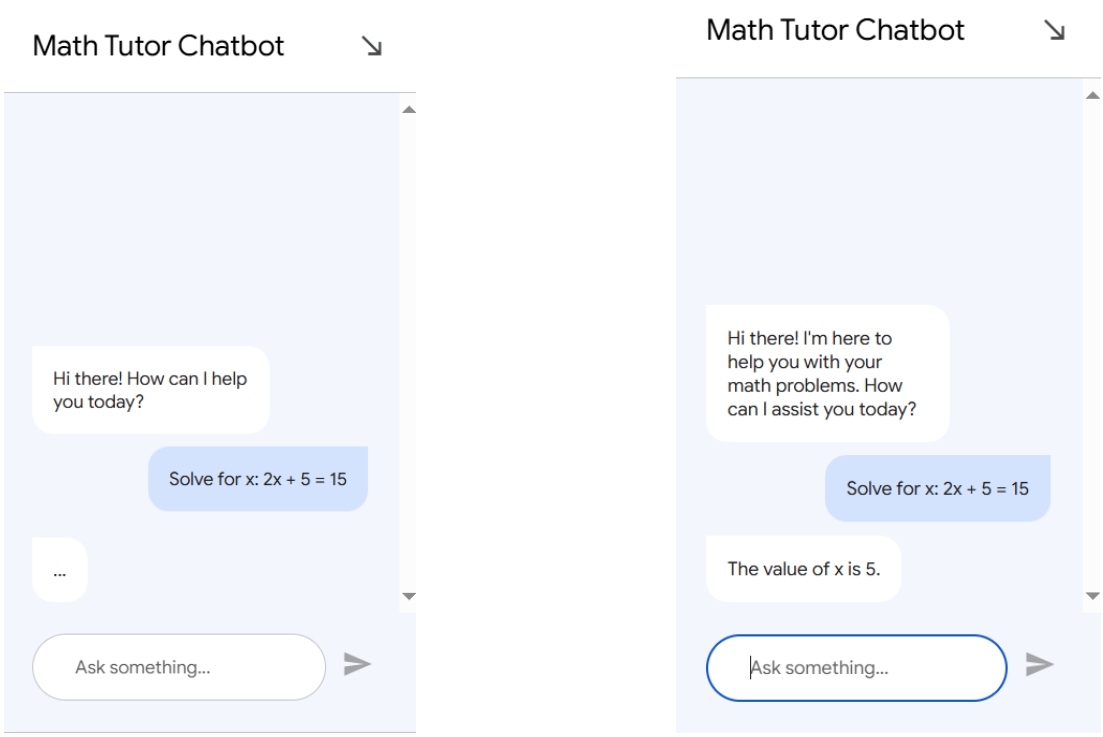

Fig 5.8.2 Solution from chatbot

5.9 Discussion

The AI-powered math tutor chatbot, built using Dialogflow, revolutionizes the way students approach learning mathematics by providing instant, on-demand assistance through conversational interactions. This system utilizes Dialogflow's powerful natural language processing (NLP) capabilities to understand and interpret students' questions in real time, offering step-by-step solutions to a broad range of mathematical problems. Whether it's solving algebraic equations, performing complex calculus operations, or helping with geometry, the chatbot is designed to handle various levels of difficulty, tailoring its responses to the needs of the individual user. The chatbot operates by recognizing intents within user input, and based on these recognized intents, it triggers appropriate responses, which are then delivered in a way that is both clear and educational.

The project makes extensive use of Dialogflow's intent recognition and webhook fulfillment to provide personalized, accurate, and contextually relevant responses. Through the integration of these components, the chatbot is able to guide students through complex problem-solving processes, ensuring that each explanation is comprehensive and easy to follow. In addition, the system continuously refines its performance through user interactions, collecting feedback on the accuracy and clarity of its solutions. This feedback is used to enhance the bot's knowledge base and adjust its algorithms for better accuracy over time.

Moreover, the system is built with scalability in mind. As new features are added, such as support for different types of math problems, additional educational content, or even multi-language capabilities, the architecture allows for seamless updates without compromising user experience. By leveraging Dialogflow's robust API and webhook features, the chatbot can also integrate with external resources, such as math-solving libraries or databases, to provide richer, more detailed solutions.

While the chatbot's performance is primarily evaluated through user feedback, accuracy, and response time, further refinements are focused on improving its ability to handle edge cases and provide even more personalized learning experiences. Through continuous optimization, the AI-powered math tutor chatbot promises to be an invaluable tool in modernizing education, providing students with reliable, accessible, and interactive math help wherever and whenever they need it.

Chapter- 6

CONCLUSION & FUTURE RESEARCH

6.1. Conclusion

This project has successfully designed, implemented, and evaluated an AI-powered chatbot for assisting students in solving mathematical problems. Utilizing Google AI Agent Builder and the Gemini generative approach, the chatbot demonstrates advanced natural language processing (NLP) and multi-step reasoning capabilities. These technologies enable the system to comprehend a wide range of mathematical queries, interpret complex problem structures, and generate accurate, step-by-step solutions with detailed explanations.

The evaluation of the chatbot's performance, based on key metrics such as response accuracy, problem-solving effectiveness, and explanation clarity, confirms its ability to address various mathematical challenges. The chatbot consistently provided relevant solutions and interactive feedback, making it a valuable educational tool. However, some limitations were observed in handling highly abstract or unconventional problems, emphasizing the need for further refinement in reasoning accuracy and contextual understanding.

Despite these challenges, the project highlights the significant potential of AI-powered chatbots in revolutionizing mathematics education. By delivering instant feedback, personalized learning support, and scalable tutoring, the chatbot empowers students to learn at their own pace. This accessibility is particularly beneficial for learners who require additional practice or individual guidance beyond traditional classroom settings.

Moreover, the integration of AI-driven solutions in education presents an opportunity to enhance conventional teaching methods. The chatbot's interactive approach makes mathematical learning more engaging, helping students develop confidence in problem-solving. Additionally, its scalability ensures wider accessibility, bridging learning gaps and catering to diverse educational needs.

However, as AI adoption in education expands, addressing critical concerns such as data privacy, fairness, and transparency is essential. Ensuring that AI-driven learning tools operate ethically and securely will be crucial in maintaining trust and maximizing their positive impact.

6.2 Future Implications

This project lays the foundation for further advancements in AI-driven education. Future improvements could focus on enhancing reasoning depth, multi-modal learning (graphs,

diagrams, voice explanations), and real-world problem applications. By refining these aspects, AI-powered chatbots can evolve into indispensable tools for personalized, scalable, and adaptive learning experiences.

In conclusion, this project demonstrates that AI chatbots built using Gemini's generative approach have the potential to redefine mathematics education. With ongoing development and responsible AI implementation, these intelligent systems can create a more accessible, interactive, and effective learning environment for students worldwide.

References

[1] Vertex AI Agent Builder and DialogFlow, "Enhancing Conversational AI with Dialogflow and Vertex AI: A Deep Dive into Vertex AI Agent Builder," 2025.

[2] J. &. B. A. Smith, "AI in Education: Revolutionizing the Classroom. Journal of Educational Technology," 2020.

[3] H. &. Z. Y. Lee, "Mathematics Education and AI-Powered Tools," *Journal of EdTech Innovation,* vol. 6, no. 1, pp. 24-38, 2022.

[4] T. &. L. P. Jacobs, "Scalable AI for Personalized Learning in EdTech," *Journal of Digital Learning,* vol. 22, no. 6, pp. 54-63, 2022.

[5] F. Raza, "AI in Education: Personalized Learning and Adaptive Assessment," 2023.

[6] A. &. W. P. Smith, "AI Chatbots in Education: Enhancing Mathematics Learning," *Journal of Educational Innovations,,* vol. 8, no. 3, pp. 34-49, 2023.

[7] S. &. K. R. Patel, "The Role of AI in Mathematics Education: Trends and Challenges," *International Journal of AI in Education,* vol. 5, no. 2, pp. 18-29, 2022.

[8] H. &. L. Y. Zhang, "Challenges in Implementing AI for Math Problem Solving," *Journal of Technology in Education,* vol. 12, no. 1, pp. 11-22, 2021.

[9] B. U. (. Zaman, "Exploring Mathematics Education with AI-Enhanced Chat Bots," ResearchGate, 2023.

[10] J. &. M. K. Williams, "Scalable AI Solutions for Personalized Math Education," *Journal of EdTech Advances,* vol. 17, no. 4, pp. 50-53, 2023.

[11] T. &. D. C. Lee, "Contextual Understanding in AI-Powered Math Tutoring Systems," *Journal of Artificial Intelligence and Learning,* vol. 3, no. 5, pp. 13-25, 2022.

[12] A. A. M. &. A. S. Alshammari, Role of AI chatbots in education: Systematic literature review. Education and Information Technologies, 2023.

[13] N. K. D. Z. G. B. E. &. Y. B. Güler, "Evaluating ChatGPT's academic achievement in a mathematics exam," Pedagogical Research, 2024.

[14] S. N. R. R. F. J. &. C.-d.-l.-S. M. Moral-Sánchez, "Analysis of artificial intelligence chatbots and satisfaction for learning in mathematics education," *International Journal of Educational Research and Innovation,* vol. 20, pp. 1-14, 2023.

[15] Y. &. C. Z. Huang, "Integrating AI chatbots into mathematics instruction: Opportunities and challenges," *International Journal of STEM Education,* vol. 9, no. 1, pp. 1-15, 2022.

[16] L. &. D. J. Carter, "Ethical Considerations in the Use of AI in Education," *Journal of Ethics in Technology*, vol. 11, no. 1, pp. 8-21, 2023.

[17] L. &. W. X. Jiang, "Designing AI-powered math tutoring systems: A user-centered approach. Computers & Education," p. 167, 2021.