VIETNAM NATIONAL UNIVERSITY OF HO CHI MINH CITY
THE INTERNATIONAL UNIVERSITY
SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

AiMA - An AI-Based Mobile System to Assist College Students with Math-Related Issues

By
DANG NGUYEN NAM ANH
Student ID: ITITIU19076

A thesis submitted to the School of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
Bachelor of Science in Computer Science

Ho Chi Minh City, Vietnam
2024

AiMA - An AI-Based Mobile System to Assist College Students with Math-Related Issues

APPROVED BY:

_____,
Le Duy Tan, PhD

THESIS COMMITTEE

Assoc. Prof. Vo Thi Luu Phuong,

Dr. Le Duy Tan,

Dr. Le Hai Duong,

Dr. Nguyen Thi Thanh Sang,

Acknowledgement

I want to express my sincere gratitude to the International University, Ho Chi Minh City (HCMIU), and our principal Tien Khoa Tran for providing me with the opportunity to complete this thesis on the topic of "AiMA - An AI-Based Mobile System to Assist College Students with Math-Related Issues". They also assisted me in conducting extensive research, and I am very appreciative of that.

I am honestly grateful for all the knowledge and help of everyone who gave me to improve and develop my thesis project day by day. I want to express my sincere gratitude to my instructor, Dr. Duy Tan Le for supporting me in entering academic competitions as well as providing the chance for me to be nearer to my research work. Since then, I have gained more confidence in continuing to develop this project. Besides, I want to send thanks to my friend Binh Nguyen Le Nguyen for suggesting many ideas to make this project unique.

I would also like to extend my sincere appreciation to my family for their unwavering support throughout the development of my thesis. Supportive of my mental health throughout my journey, my family was always by my side.
January,
Dang Nguyen Nam Anh.

Contents

List of Figures

List of Tables

LIST OF ABBREVIATIONS

UI **U**ser **I**nterface

UX **U**ser **E**xperience

TMC **T**heoretical **M**odels in **C**omputing

RTK **R**unge **K**utta

ODE **O**rdinary **D**ifferential **E**quations

MVC **M**odel **V**iew **C**ontroller

URL **U**niform **R**esource **L**ocator

ID **I**dentification

OCR **O**ptical **C**haracter **R**ecognition

CNN **C**onvolution **N**eutral **N**etwork

ViT **Vi**sion **T**ransformer

Hybrid-ViT Hybrid **Vi**sion **T**ransformer

ReLU **R**ectified **L**inear **U**nit

GPU **G**raphics **P**rocessing **U**nit

Abstract

As technology continues to progress, the importance and influence of mathematics in people's lives are growing. The emergence of mobile systems has brought about several benefits, such as portability, convenience, and accessibility, especially in the field of education. It is crucial to develop a mobile system that can cater not only to high school students but also to those enrolled in colleges and universities, offering solutions for basic calculations and highly specialized algorithms. To address this need, this research presents AiMA, an AI-based mobile system that assists college students with math-related issues. In this paper, we discuss the limitations of current math-solving mobile applications and propose a solution that caters to specialized algorithms for university students. AiMA employs React Native on the front-end side to render the user interface and Flask-server on the back-end side to handle numerical data. Additionally, Hybrid-ViT - an AI architecture for OCR tasks is implemented to handle input images stored in a cloud-based media management platform for detection. The system's functionality evaluation indicates that AiMA provides advantages over existing solutions by enabling users to solve complex calculus problems and receive tailored solutions for specific areas of mathematics. This makes AiMA a highly valuable tool for college and university students who require specialized assistance with their math-related studies.
Keywords: optical characters recognition, mathematics, mobile application, deep Learning, latex.

Chapter 1

OVERVIEW

1.1 Introduction

1.1.1 Mobile application background

A mobile application ("app") is a term that is often used to describe programs that run on mobile devices such as smartphones and tablets [1]. The major distinction between regular desktops and mobile apps is that mobile applications may be obtained via the app store. When applications are downloaded, most devices instantly install them, making the installation procedure simple for the user. Furthermore, since they must be more powerful and take less resources to produce, most mobile apps are low-cost. However, because of the restrictions of mobile operating systems (such as iOS or Android), apps may provide fewer capabilities than desktop applications. However, since its portable, small-size, easy-to-use, and quick-response properties [2], the app sector has become the trend in the creating community. As mentioned, the strongest point of the mobile platform is its convenience and streamlining. So, as of now, Mobile apps can be used to perform many different types of tasks, including:

- Communication: most mobile applications today, especially social networking applications, always allow users to communicate with each other by text messages and phone or video calls [3]. Thanks to these applications, the problem of distance in communication is no longer an obstacle, we can talk or meet each other anywhere and at any time.

- Entertainment: most users tend to entertain when using mobile phones, so entertainment applications such as video games, movies, or music-listening applications are always top choices [4].

- Information: The need to monitor information is a necessary need, users will often look to electronic press applications, social networks, or applications that synthesize air quality-monitoring information to understand capture, and update information.

- And especially in the field of education, with mobile applications, learners have more ways to interact with the learning environment than ever before. This is also a new but extremely effective learning method. It promotes students' self-study ability, making learning more convenient and interesting [5]. These include online learning applications that can provide answers and solutions, in addition, to creating a learning community and helping learners to discuss problems in class together [6].

It can be said that this is a big step forward in education as digital transformation can now be integrated to optimize learning.

From there, we can see that mobile applications are becoming an indispensable part of people's lives. The strong growth of mobile applications is due to a combination of many factors, including the popularity of smartphones as today, it has become an essential device in people's lives. According to Statista statistics, as of January 2022, there are about 6.4 billion smartphone users worldwide [7]. This creates a huge potential market for mobile applications. In addition, the development of technology is also the reason why mobile platforms have become popular. Thanks to rapidly developing technology, it brings many new features to mobile applications. Mobile applications today can perform many complex tasks, from shopping and entertainment to studying and working as mentioned above. And especially user needs, users increasingly need to use mobile applications to meet their needs quickly, conveniently, and anytime, anywhere. Mobile applications provide numerous advantages to both individuals and enterprises [8]. Mobile applications provide users with convenience, and enjoyment, and improve their quality of life. Mobile applications are a great way for businesses to reach clients, assisting in the rise of sales and the enhancement of brands. Mobile applications will continue to evolve and become an integral component of people's lives in the future. Mobile apps will become smarter, more personalized, and integrated with a wide range of devices.
Mobile applications are now deployed on multiple platforms, the most popular are iOS and Android [9], and the latest one is the super app platform of Zalo [10]. Each platform has its strengths, such as iOS with its easy-to-use interface and security. Applications running on iOS are particularly good when running on Apple devices with powerful configurations and can optimize performance. In return, applications running on this operating system have low compatibility, most of them will not work properly on other platforms [9]. As for the Android platform, multitasking is a feature of Android application development. As a result, a user can use two or more applications at the same time. This allows streaming through social accounts and song listening. Besides, open-source applications are its strength, as there are many free open-source applications available in the Android environment. Thousands of apps are available for free in the Android marketplace [9]. In addition, framework and library are also two important concepts in mobile app development. frameworks help developers save a lot of time and effort during the development process by reusing built-in functions. A library is a collection of pre-built functions and classes, that help programmers perform tasks more easily. Popular frameworks in mobile application development include Flutter, a framework that supports cross-platform application development. Flutter uses the Dart language to create interfaces and features for applications. Flutter can reuse code to deploy on many different platforms and operating systems such as iOS and Android, as well as run on the Web App platform [11]. Besides Flutter, React Native is also a powerful framework, developed by Facebook. React Native is written in JavaScript language. React Native uses native components to create user interfaces, which allows programmers to organize their code better. In addition, React Native can support many operating systems such as iOS and Android [12]. This saves more development time and effort when developing cross-platform.

1.1.2 Mathematics background

Mathematics is an abstract subject, but it has extremely wide applications in our lives. It can be considered an essential tool for helping people better comprehend their surroundings and solve problems. Furthermore, in the current era of advanced technology, when

everything is automated and "smarter", mathematics is asserting its function more and more, exerting a substantial influence on several subjects, particularly technical ones like computer science, space engineering, electrical engineering, etc [13].
Especially, mathematics is crucial in computer science, particularly in the domains of deep and machine learning [14]. Mathematics is employed in deep learning to formulate, optimize, and assess models. Machine learning models frequently employ intricate mathematical algorithms, such as optimization algorithms for determining optimal parameter values [15], probabilistic algorithms for computing outcome probabilities, and statistical algorithms for data analysis. Subsequently, one can execute distinct operations like image and text classification; data prediction, and automation functions. The present-day artificial intelligence models depend heavily on mathematical formulas. Mathematics is utilized in image processing to create algorithms that enable facial identification, handwriting recognition, and classification of images (14). In the domain of natural language processing, mathematics is employed to create automated translation algorithms, voice recognition systems, and emotion analysis techniques.
Moreover, mathematics is utilized in various domains. for instance, in the field of finance, mathematics is employed to create automated trading algorithms based on technical analysis and statistical analysis, analyze the stock market, and implement risk management strategies [16]. Math is applied in the field of medicine to create algorithms for disease detection and genetic research [17]. In the industrial sector, this subject is employed to develop and implement automation systems, enhance manufacturing efficiency, and monitor supply chain operations. Therefore, the importance of mathematics in human life and development cannot be denied. Mathematics is not only a useful tool in the fields of science and technology but also an important factor in discovering a deep understanding of the world around us. From solving everyday problems to shaping the future through research and innovation, mathematics plays an irreplaceable role in the advancement of society. In short, the importance of mathematics is in the academic field and spreads to all aspects of life, positively contributing to the comprehensive development of people and society.
The impact of this subject on social life has been seen. For students who regularly have access to math, what impact will mathematics have on them? Undoubtedly, mathematics holds an important place as an academic field, exerting a profound influence on students. Mathematics not only builds students' understanding of arithmetic, algebra, geometry, and other mathematical concepts, but also has a beneficial influence on the enhancement of their mental skills, intellect, and proficiencies. Mathematics facilitates the development of logical reasoning and the acquisition of problem-solving skills in students. Mathematics equips pupils with crucial abilities, such as logical thinking, critical examination [18], and the ability to combine information into a coherent whole. These abilities enable students to effectively address challenges encountered in their daily lives, such as managing finances, strategizing, and making informed choices. When students tackle a ratio problem, they must employ logical thinking to determine the correlation between two numbers or objects. Students must engage in the analysis of information to select the most advantageous solution. To make informed decisions, students must synthesize the information and arrive at the most logical and persuasive conclusion.

1.1.3 Optical Character Recognition Background

Optical character recognition (OCR) is a branch of artificial intelligence (AI) that focuses on the automated conversion of text presented in images into text that can be understood

by machines [19]. This is a computer technology that converts text in images into text readable by computers. OCR can be used for a variety of purposes, including digitizing documents, OCR can be used to convert paper documents into electronic documents, making them easier to store and access. Data entry automation, specifically this technology can support automatic data entry, such as entering information from invoices or credit cards with just their images [20]. Additionally, OCR can also be used to search for information in digitized documents. This technology, because it interacts completely with the image, so its accuracy and how it works will depend on the clarity of the image to be recognized. Therefore, OCR always has a step of pre-processing the data, which includes removing noise, improving contrast, and reformatting the image to fit the OCR algorithm [21]. After that, the model was able to identify the characters in the image and from there proceed to identify and return prediction results.

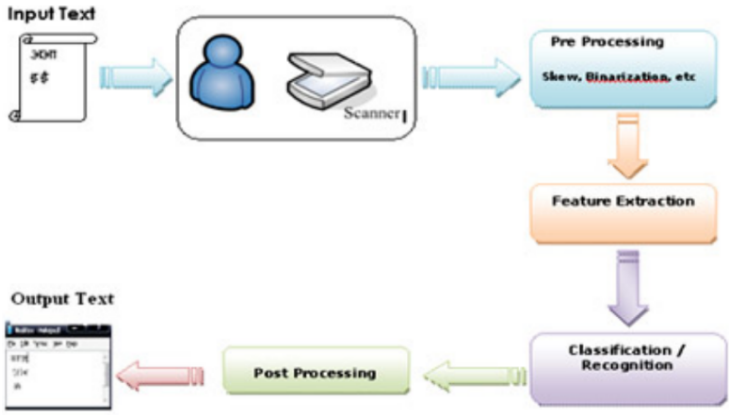


Figure 1.1: OCR process diagram

Optical character recognition (OCR) technology has brought many opportunities for advancement in the field of education. By converting text from images or documents into processable data form, OCR has made possible many important applications. According to Gourab Nath in the article "Isolated OCR For Handwritten Forms: An Application in the Education Domain", the applications of OCR in this field can be mentioned as the ability to systematize the learning process [22]. First, textbooks can be scanned and converted into electronic format, making it easier for students and teachers to access and find information. In addition, OCR can also be combined with other NLP models, helping students with poor reading ability to be supported through text-to-speech conversion, opening up broader educational opportunities for them. Furthermore, organizing and storing research documents becomes more convenient. OCR helps create searchable documents, reducing the time spent searching for information. During the teaching process, recognizing text from videos and converting images from whiteboards helps support online teaching processes. Besides, integrating OCR into the exam marking process also brings significant efficiency, reducing pressure on teachers and increasing fairness in the evaluation process. Automating the exam marking and essay evaluation process saves time and increases accuracy. Most recently, OCR technology has also been integrated into applications that support student learning such as math applications. This technology is used to scan mathematical equations, which is especially convenient, users do not need to type the equation to be solved using the keyboard, which makes the user experience worse because it is time-consuming. Now, users only need to use the camera to scan the image of the equation to be calculated, and the system will return results quickly and accurately. This is considered an improvement, making math applications

more convenient by optimizing the user interface.

Figure 1.2: OCR of latex conversion for mathematics' equations recognition

As in the above image, each mathematics equation or formula can be displayed by a normal string following the convention of latex. Hence, the lengthy and complex problems can make the model more challenging to recognize.

1.2 Problem Statement

Overall, mathematics is an important subject that has a positive impact on students. Studying math well will help students develop comprehensively in thinking, and intelligence and prepare for the future. Nevertheless, because of the high difficulty of mathematics, students always face enormous challenges in studying. Because this is a subject that requires flexibility, creativity, and a solid knowledge base, many students find this subject very difficult to absorb and apply. Specifically, mathematics covers many fields, such as calculus, linear algebra, and geometry. Mastering mathematical concepts, formulas, and rules requires a long and patient learning process. Moreover, as the above mention, mathematics requires students to have the ability to think logically, reason, analyze, and synthesize. Many students have difficulty solving mathematical problems, especially difficult or highly applicable problems. This leads to some students finding math dry, confusing, and boring. This makes them lose interest in learning and not want to learn math. There will even be a part of students who feel scared when learning this subject. According to Ishfaq AHMAD Bhat's research [23], many students fear this subject like they fear evil. They feel math is not important and will lower their academic scores. And learning math is just chasing scores, in other words, they are being forced to learn this subject. Besides, following the research of K. Abdul Gafoor's team [24] from Farook Training College, Kozhikode, Kerala, they stated that there is a random sample of 51 standard IX students who were given a questionnaire to fill out. The questions asked about their expectations in math, their beliefs regarding task worth, their self-efficacy, their beliefs about how we know what we know, their interest, and their anxiety. Of these students, 88% said mathematics was their most disliked subject, while only 6% said it was their top choice. The most common causes for disliking mathematics were difficulties in understanding the subject content and teacher or instructional-related concerns. Mathematics was ranked as a very difficult subject by 20%, a medium difficulty topic by 54%, and an easy subject by 10%. And 42% of students are unable to discover solutions to issues presented in their textbooks. In addition, although the traditional teaching method is always the most effective, in the context of social development, where everything is

moving towards convenience and speed, sometimes learners will still feel uncomfortable and find it more difficult to absorb traditional teaching methods. Because we know that each student has a different ability to receive knowledge. Combined with a somewhat passive learning method, it unintentionally leads to an imbalance in the abilities of each student in the learning environment.
Due to social changes in technology, many new, modern, convenient, and economical tools for learning and education have appeared today. There are mobile applications that can support students in the learning process. Currently, the above types of applications have appeared quite a lot with many different variations and forms. Such as dictionary applications, timetable management applications, or applications that support English learning. But perhaps the most prominent one can mention calculation applications, because they act as a home teacher, helping students solve homework effectively. However, these applications are still facing limitations, which is that most of them can only support card types within a certain limit. That is, they have difficulty solving more advanced, specialized problems. This causes users to not be fully supported, the applications they download can only solve common types. There are even still applications that can only solve basic types of math, for problems of the same type, but require more complex calculations, these applications cannot support or provide answers exactly. In addition, the application interface is also an issue worth paying attention to, as nowadays, many applications are challenging to use. Specifically, suppose a completely new user is using an application for the first time, with a confusing and confusing design that will confuse the user. Finally, in today's applications, most require users to purchase their services to activate all system features, specifically the solution display feature. This is especially inconvenient when this feature is one of the basic features of a support application. In addition, no matter how powerful and diverse the support application is, one thing is certain: the system will not be able to 100% support all student problems. Surely there will be situations that machines cannot solve. Thus, solving academic problems automatically is not enough. The plan proposed here can create an environment or forum for students to discuss difficult problems with each other. This will improve the practicality and feasibility of the application, instead of just providing users with a rough system. Although the idea of a math application is completely valid and can make a significant contribution to education, implementing it properly and following market trends is something that needs attention
The current impact of artificial intelligence on societal progress is significant. In addition to sectors such as healthcare, industry, and manufacturing, education also has major possibilities for the implementation of these contemporary advances. Mathematics presents a huge challenge for both students and individuals who encounter it in their daily professional endeavors. This field necessitates learners to possess a strong fundamental basis and also to be sensitive in integrating knowledge.

1.3 Scope and Objective

Recognizing the problems mentioned above shows that there still exist certain limitations in mathematical applications on the market, as well as AI models used for the identification of mathematical characters. Therefore, inheriting their strengths and improving their weaknesses is a wise decision.
This graduation project proposes a cross-platform mathematics application that uses OCR technology. The mobile app will be able to solve problems from common aspects such as geometry, and calculus to specialized types of problems, with more advanced and complex knowledge such as numerical methods, probability, and statistics. This application was

developed mainly to contribute to the community because research on the general level in the market shows that currently there are very few open-source mathematics applications that can fully meet the learning needs of students. Furthermore, because this application was developed to support students in their studies, it can optimize the user's self-study experience, so in addition to providing answers for math problems, the application can also provide specific methods for many various types and fields. The application also receives attention in optimizing even the smallest factors such as the interface. An easy-to-use interface and providing complete instructions for use is a wise way to help retain users. The interface is built with a focus on minimalism, which means it will be easy to see, and easy to interact, and users can completely get used to it after only 1 to 2 uses. In addition, the performance of the application is also taken care of for optimization. By applying cache-saving methods, to using framework components appropriately, the application will be able to deliver a fast user experience. As mentioned above, the strength of the application compared to other predecessor applications is the ability to solve specialized problems with high complexity, so this application will focus on building algorithms for those problems. It is expected that at least 30 advanced algorithms will be created to fully support users. This promises to be an application that contributes highly to society. However, there will certainly still be limitations that machines cannot solve. Therefore, the forum feature was created, to connect users, users can now interact to exchange with each other about pressing issues.
In short, this application can meet the criteria, from basic to the ones able to contribute to society as follows:

- Calculation ability: the application is capable of calculating all types of problems from basic to advanced which are popular with students and engineering colleagues in Vietnam, with almost absolute accuracy.
- Easy to use: the application is designed to be streamlined, with a friendly interface, easily accessible to most users.
- High performance: the application's response and calculation capabilities are always being improved to give users the fastest and most convenient experience.
- Diversifying features: besides calculation features, the application integrates other valuable features such as the forum feature to connect learners, and most importantly, the recognizing feature for the mathematical characters.
- Improved model: the OCR model in the system will be a complete version and have stable performance when it can recognize many types of equation images, including handwriting.
- Low development costs: because this project is implemented mostly by reusing open-source libraries and frameworks, the development costs will be very low.
- Contribution: This application was created to contribute an open-source mathematics application to the field of education, which currently very few people have done in the same aspect.

This is also the end of the introduction part of the report. In the next part, the report will present an overview of the advanced algorithms that the application can support, in addition, information about the technologies used in the application, from the development of technology mobile app to optical character recognition models will also be clarified.

The third part of the report will be a study of publicly available applications and models, from which the strengths and weaknesses can be evaluated, and then compared with this proposed application. The fourth part will be a detailed description of the implementation method, building the system, from initializing the interface to building the server for calculation, and the process of improving the AI model. Part 5 will be the results part, the report will describe the elements that the application has been able to perform, and from there will be discussed and evaluated in part 6. Part 7 will be the future work part, After the objective assessment, the remaining limitations will be listed for resolution shortly.

Chapter 2

LITERATURE REVIEW

2.1 Research in mathematics fields

Because this is a math application, math-related topics are always of interest for development. Mathematics covers many fields, each field has its characteristics. From basic and common fields to more special and complex aspects, all have different applications in social life. Therefore, the main research area of this thesis report will focus on 5 major topics in mathematics, which are algebra, calculus, geometry, numerical methods and probability, statistics, and random process problems. In the first part of this section, the article focuses on providing background knowledge for each of the above fields.

2.1.1 Calculus & Algebra

Since mathematics is such a wide and intricate field, it is frequently separated into numerous subfields. There are basic computations related to addition, subtraction, multiplication, division, and other number concepts like fractions, decimals, or mixed numbers covered in common areas. Moreover, we have Calculus, a more sophisticated option that covers deeper ideas like integration and differentiation.
The study of function changes is carried out through differential calculus. It enables us to determine the area under a curve or the rate of change of a function at a given point [25]. The derivative of a function can also be found mathematically using differentiation. The rate of change of a function at a place is described mathematically by the concept of the derivative. The function y = f(x) has the following differential formula: $\frac{dy}{dx} = f'(x)$
In there,

- dy is the change of the function value y at point x.
- dx is the change in variable x.
- f'(x) is the derivative of the function y at point x.

For example, the function $y = x^2$ has a derivative of $2x$. This means that when x changes by one unit, the value of the function y will change by $2x$ units. There are many applications of calculus in life, such as finding the velocity of a moving object, or the area or volume of complex shapes. With a higher order function, the derivative of that function is a higher order function. For example, the derivative formula for functions of degree, formula would be $f'(x) = n \cdot x^{(n-1)}$
As for complex problems, specifically the derivative of a function containing other functions. For example, the derivative of the function $y = \sin(x^2)$ is a function that contains

the derivative of the function $\sin(x)$ and the derivative of the function x^2. This can make calculating the derivative of the function more complicated. The derivative formula of a function containing other functions is calculated according to the following formula: $f'(x) = \frac{d}{dx} f[u(x)] = f'[u(x)].u'(x)$
In there,

- $f(x)$ is a function that contains the function $u(x)$.
- $u(x)$ is a function contained in the function $f(x)$.

The mathematical integration process synthesizes a function's values at unlimited points within a certain period. Integration is only the derivative calculated in the opposite direction [26]. Finding the velocity at a given time is calculated using the derivative, and calculating the distance traveled from a starting time to the present is calculated using the integration method. Integrals come in two primary categories: Definite integral: is a function's integral over a certain interval. For instance, $\frac{x^3}{3} \cdot (b-a)$ is the definite integral of the function $f(x) = x^2$ in the interval [a, b]. Next, an indefinite integral is an integral of a function that is not aware of its integration's limit is called an indefinite integral. For instance, the function $f(x) = x^2$ has an indefinite integral of $\frac{x^3}{3} + C$, where C is a constant. Similar to derivative, anti-derivative has many formulas, which leads to the difficulty level for the beginner when they get used to learning this.
There are many ways to solve an anti-derivative problem, which makes this type of problem more complicated when students are required to combine and apply many types of knowledge flexibly, such as finding the root function using the formula table, this is the most basic form because you only need to memorize and apply the available formulas. This type of problem is not a math puzzle. To solve this type of problem, we need to do the following steps:

- Determine the form of the function to be integrated.
- Choose the appropriate integral formula.
- Perform calculations.

For example, we have the problem of calculating the area of a flat figure limited by the curve $y = x^2$ and the horizontal axis on the segment [0, 2]. To solve this problem, we need to consider the function $f(x) = x^2$. This function is a 2nd-order exponential function, so we can apply the primitive formula of the 2nd-order exponential function, we get $\int x^2 dx = \frac{x^3}{3}$, replace x with small and large values in the given segment, we get the area we need to find is $\frac{8}{3}$.
The above-mentioned above is just a part of anti-derivative problems, there is another method called partial integral. This method has the advantage that it can be applied to complex functions. In addition, this method can also be used to solve integral problems, such as the problem of calculating area, volume, etc. Specifically, we can apply this to solve problems containing exponential functions and trigonometric functions such as exponential functions, trigonometric functions, composite functions, and fractional functions. This method is based on the following formula: $\int udv = uv - \int vdu$
In which, $u(x)$ and $v(x)$ are two functions with continuous derivatives on K. To use the partial primitive method, we need to perform the following steps:

- Determine the components $u(x)$ and dv in the function.
- Calculating the derivative $u(x)$, we find du.

- Calculate the integral function dv, we have v.
- Apply the above formula to calculate the integral function.

For example with the problem, we have: $\int x \ln x dx$. First, we define $u = lnx, dv = xdx$. Then, we can find du by calculating the derivative of $\ln x$, applying the formula, we have $du = \frac{1}{x}dx$. Next, we can calculate the antiderivative of dv to get v, so we have $v = \frac{x^2}{2}$. Finally, applying the above formula, we get the result for this problem as $\frac{x^2 \ln x}{2} - \frac{x^2}{4} + C$.
There are also many related types of problems, which shows that calculus is a field with many applications in life, however, it is also a quite challenging subject for middle school and high school students and university students.
Besides differentiation and integration, there are many other problems related to mathematics, such as calculating the limits of functions and equations with degrees. Talking about equations with a degree, which is related to the algebra aspect, this is an equation whose highest term has degree n, for example, $2x^n$. Specifically, equations with degree 1 are called first-order equations, and equations with degree 2 are called second-order equations, similar to other higher-order equations. An equation of degree n can be written as: $a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 = 0$, where, $a_n \neq 0$ and n is an integer. This field concludes with many types. For each level of the equation, there will be a separate solution, such as for a quadratic equation, this is an equation of the form $ax^2 + bx + c = 0$, where a, b, and c are the real numbers and $a \neq 0$. This method is general, and applicable to all quadratic equations. In addition, this method is simple, easy to remember, and easy to apply. The number of roots of a quadratic equation can be determined by the discriminant formula: $\Delta = b^2 - 4ac$
The number of values of x will be determined by comparing the delta with zero, in detail, with a positive delta, the quadratic equation has two solutions, while the delta equals zero, the equation has only one solution, and finally, there are not any solutions with negative delta.
Specifically, with a delta greater than 0, we will apply the formula below to calculate the two solutions of the equation. For example, we have $x^2 + 2x - 3 = 0$. First, we will calculate $\Delta = b^2 - 4ac = 2^2 - 4 \times 1 \times (-3) = 20$. We see, $\Delta > 0$, so the equation has two distinct solutions. We can apply the discriminant formula to calculate the two solutions of the equation: $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{(-2 \pm \sqrt{20})}{2}$.
Besides quadratic equations, we have a solution for fourth-degree equations. The method of solving this equation by setting $t = x^2$ is a quite simple and effective solution method. This method is based on converting a quadratic equation to a quadratic equation. Regarding how to solve this problem, we can do the following:

- First, set $t = x^2$.
- Then replace x^2 with t in the 4th-degree equation.
- Quadratic equation currently can be solved.
- Finally, solve the equation $x^2 = t$ to find the solution of the original quadratic equation.

For example, with the problem $x^4 - 4x^2 + 3 = 0$, first, we will replace $x^2 = t$, rewrite the above equation, we have $t^2 - 4t + 3 = 0$. Apply the discriminant formula mentioned above for this equation, we obtain 2 solutions, 1 and 3. Finally, we solve the equation $x^2 = t$ to find the solution of the original quadratic equation, we get $x = \pm 1$ and $x = \pm \sqrt{3}$.

Another method is the method of factoring equations. This method is based on Viet's theorem, according to which the sum of the roots of a quadratic equation is $-b/a$ and the product of the roots is c/a. The idea of this form of math is as follows:

- Find two numbers x_1 and x_2 such that $x_1 + x_2 = \frac{-b}{a}$ and $x_1 x_2 = \frac{c}{a}$.
- Factor the equation with the factors $x - x_1$ and $x - x_2$.
- Solve the obtained equation.

In conclusion, the ideas for calculus and algebra solving do not stop at this point. There are too many types and forms of exercise with various methods to solve. In short, Calculus is an important subject in mathematics, widely applied in many fields such as physics, engineering, economics, etc. Calculus problems often require calculating extreme values, integrals, derivatives, etc. of functions. To solve these problems, you need to master the basic knowledge of calculus, including concepts, theorems, and calculation methods. Algebra is an important subject in mathematics, widely applied in many fields such as linear algebra, analysis, arithmetic, etc. Algebra problems often require finding solutions to equations and inequalities, systems of equations, etc. Although these are popular aspects, they are very challenging for students because they require a lot of knowledge and the ability to apply knowledge flexibly.

2.1.2 Geometry

Basic Geometry

Geometry is a branch of mathematics that deals with questions of shape, size, relative positions of shapes, and properties of space. Geometry includes formulas for length, area, and volume, such as the Pythagorean theorem, circumference and area of circles, area of triangles, and volumes of cylinders, spheres, and pyramids [27]. The basic geometric shapes mentioned above all have formulas to calculate and describe the characteristics of shapes, such as perimeter and area. These are two important concepts related to geometric objects and are widely used in many contexts.

Area

Area is a concept that measures the "surface" of a geometry. For flat geometries such as triangles, squares, or polygons, the area is often calculated using special formulas corresponding to each shape. For example, the area of a square can be calculated by multiplying its length and width.

Perimeter

Perimeter is the measurement of the "boundary" of a geometric figure, which is the sum of the lengths of all the sides of the figure. Just like area, the formula for calculating perimeter depends on the specific type of geometry. For example, the circumference of a circle can be calculated using the formula $C = 2\pi r$, where r is the radius of the circle.

2.1.3 Numerical methods

Numerical methods are mathematical methods that are based on the use of numerical techniques. Numerical approaches are used to resolve mathematical issues in many domains that are unable to be solved by traditional mathematical equations or classical methods [28]. Instead, this approach employs numerical and computational techniques. In simple terms, this method uses algorithms to solve approximate, highly complex problems that are difficult or even impossible to solve using calculus and algebra methods. The following problems are addressed using techniques:

Optimize methods

Optimization is a branch of applied mathematics, concerned with the development and use of algorithms to find the optimal value of a function [29]. The optimal value can be a minimum value, a maximum value, or a near-optimal value. Optimization methods have wide applications and play a significant role in many aspects of social life, especially in engineering aspects. It can be used in business and industry, medicine, and scientific research. In business, the optimization method is applied to optimize the manufacturing process, supply chain, and marketing strategy to earn the highest profit. In industry, it optimizes the process to avoid the waste cost in production. Applications also extend to engineering and technology, where optimization is used in electronic circuit design, computer network operations management, and many other fields. In medicine, it helps optimize treatment schedules and choose the right medication. In the field of energy and environment, optimization methods are applied to reduce emissions and optimize resource use. Additionally, in the field of machine learning and artificial intelligence, optimization is used to tune model parameters to optimize prediction performance. There are many different types of optimization methods, depending on the type of function to be optimized and the constraints of the problem. Some common types of optimization methods include:

- Golden Section method: The Golden Section technique is an efficient search algorithm used to identify the precise location of the smallest point of a continuous function within a certain interval [30]. This approach relies on dividing the search range into proportional segments and retaining the segment with the most favorable characteristics. The Golden Section approach is often favored due to its avoidance of derivative calculation and its stability when applied to continuous functions. Nevertheless, this approach may need many iterations to compute the function's value, hence augmenting the computing expense for intricate functions. The Golden Section approach has many advantages, such as its straightforwardness, avoidance of derivative calculations, and capability to handle functions with discontinuities or extreme points. Hence, for intricate equations that pose challenges or are impractical to differentiate, the golden section approach serves as a fully feasible substitute. For details, to calculate a golden section problem, we can apply the following method:

  Firstly, the golden ratio can be displayed by $R = \frac{\sqrt{5}-1}{2}$

  For each iteration, it can be calculated: $d = R \times (x_u - x_i)$, then $x_1 = x_l + d$ and $x_2 = x_u - d$

  Calculate the interval $= x_l - x_u$, then check, if $f(x_1) > f(x_2) : x_l = x_2, x_{otp} = x_2$, otherwise, $x_u = x_1, x_{otp} = x_1$

  Calculate the error: $\epsilon_a = (1 - R) \times \frac{\text{interval}}{x_{otp}} \times 100$ and stop when $\epsilon_a < \epsilon_s$. Repeat from step 1

- Parabolic interpolation: this is an interpolation method that uses a quadratic polynomial to approximate values between known data points [31]. In parabolic interpolation, we assume that the interpolated function has the form of a quadratic polynomial (parabolic polynomial). Parabolic interpolation is often used when you have a set of data points and want to build a parabolic curve (quadratic function) through those points to predict values at the midpoints. Common cases include:
  - Predict values between known points: When you have several data points and want to estimate the value between two points.
  - Modeling uneven data: When data points are uneven and you want to model them with a smooth curve, parabolic interpolation can be used.
  - Estimate the peak of a graph peak: When you want to determine the peak of a quadratic function graph based on several known data points.

  In addition, there is a very interesting application of this algorithm, which is the automatic image generation feature of Photoshop. Suppose a cropped photo loses part of the image's data, then with the latest version of Photoshop, it can predict those missing image pixels, helping to refill the photo. This is the application of the interpolation algorithm, using the pixel points on the original image, the algorithm will be used to calculate the prediction for the remaining missing pixels. This algorithm can be implemented as follows:

$$x_3 = \frac{f(x_0)(x_1^2 - x_2^2) + f(x_1)(x_2^2 - x_0^2) + f(x_2)(x_0^2 - x_1^2)}{2f(x_0)(x_1 - x_2) + 2f(x_1)(x_2 - x_0) + 2f(x_2)(x_0 - x_1)}$$

  First approach, update $x_0, x_1, x_2 = x_1, x_2, x_3$, respectively. Second approach, if $f(x_1) < f(x_3)$: and if $x_1 < x_3$, update $x_0 = x_1, x_1 = x_3$, otherwise, $x_2 = x_1, x_1 = x_3$. In case of $f(x_1) >= f(x_3)$: if $x_1 < x_3$, update $x_2 = x_3$, otherwise, update $x_0 = x_3$

- The Newton-Raphson method, frequently referred to as the Newton method, is a highly effective numerical methodology for solving equations. The Newton-Raphson Method is a process that requires iteratively improving an initial estimate to approach the target solution [32]. Newton's Method is applied to estimate the roots or zeros of real-valued functions. It involves starting with an initial guess (x_0) and then iteratively estimating the next guess (x_1) which is closer to the actual roots. This is done using the following.

  It can be implemented by: At each iteration, update $x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$, then check $\epsilon_a < \epsilon_s$

  And about its applications:
  - Solving non-linear equations: Newton's method is often used to find solutions to non-linear equations and systems of non-linear equations.
  - Image processing: In image processing and computer vision, the Newton method can be used to improve image quality, especially in problems such as image synchronization or image restoration.

Differential equation

In numerical methods, ODE, which stands for Ordinary Differential Equation) is an important tool used for modelizing the continuous system by time. ODE relates to the unknown function and its derivative, which describes the relationship between the change of the dependent variable (often time) and one or more other independent variables [33]. Some common numerical methods used to solve ordinary differential equations include:

- Euler method, which is the simplest method. It uses the derivative approximation formula to update the value of the function at each time step [34]. Moreover, it can be easily implemented but has low accuracy for some problems. We can apply the following formula: $y_i + 1 = y_i + f(x_i, y_i)xh$, with h as step size, and the Euler method can be looped with $\frac{|x_1 - x_2|}{h}$ times. Euler's method, although simple, has many important applications in various fields. In dynamical system simulation, this method is often used to determine the behavior of mechanical, physical, or even biological systems. It is also applied in the problem of sizing materials to estimate their changes over time, especially based on temperature or other factors.

- Runge-Kutta method: Includes many variations such as RK3 (3rd order Runge-Kutta), RK4 (4th order Runge-Kutta) method, Classic fourth-order Runge-Kutta method, etc [35]. It provides higher accuracy than the Euler method and is widely used in many practical applications, such as physics, engineering, biology, finance, etc. Especially, 4th order Runge-Kutta:

  - Financial Calculation: In the field of finance, Runge-Kutta is often used to solve differential equations related to models of stock prices and other financial instruments.
  - Circuit simulation: If you are considering a complex circuit, the Runge-Kutta method can help simulate fluctuations in current and voltage over time, helping to specify the influence of factors such as resistors, capacitors, and coils [**nyandieka2023chaotic**].

  To solve this problem, with RK3:

Third-order RTK

$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 4k_2 + k_3)h$$

- $k_1 = f(x_i, y_i)$
- $k_2 = f\left(x_i + \frac{1}{2}h, y_i + \frac{1}{2}k_1 h\right)$
- $k_3 = f(x_i + h, y_i - k_1 h + 2k_2 h)$

Fourth-order RTK

$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)h$$

- $k_1 = f(x_i, y_i)$
- $k_2 = f\left(x_i + \frac{1}{2}h, y_i + \frac{1}{2}k_1 h\right)$
- $k_3 = f\left(x_i + \frac{1}{2}h, y_i + \frac{1}{2}k_2 h\right)$
- $k_4 = f(x_i + h, y_i + k_3 h)$

Classic Fourth-order RTK

$$y_{i+1} = y_i + h\left(\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4\right)$$

$$k_1 = f(x_i, y_i)$$

$$k_2 = f\left(x_i + \frac{1}{2}h_1 y_i + \frac{1}{2}hk_1\right)$$

•

$$k_3 = f\left(x_i + \frac{1}{2}h_1 y_i + \frac{1}{2}hk_2\right)$$

$$k_4 = f(x_i + h_1 y_i + hk_3)$$

Figure 2.1: Third RTK, Fourth RTK, Classic Fourth RTK formulas respectively

Differential and Integration

- Trapezoidal Rule: commonly referred to as the trapezoid rule or trapezium rule, is a method used for numerical integration. The trapezoidal rule operates by estimating the area under the graph of the function f(x) as a trapezoid and determining its magnitude [36]. With this formula $I = (b - a)\frac{f(a)+f(b)}{2}$. The trapezoidal method is

often preferred when better accuracy than the Euler method is required and when method stability is important. For some problems, especially when studying the dynamics of dynamical systems, the trapezoidal method is a popular choice.

- Simpson's rule: Simpson's rule is a numerical method employed for the evaluation of definite integrals. Typically, the definite integral is determined by employing the fundamental theorem of calculus, which necessitates the application of anti-derivative integration techniques [37]. Simpson's rule has two types, the first is Simpson 1/3 and the next one is Simpson 3/8. Each type has its way of implementation:

  With Simpson 1/3 rule:

$$I = \frac{h}{3}[f(x_0) + 4f(x_1) + f(x_2)], h = \frac{b-a}{2} = \frac{x_2 - x_0}{2}$$

  While Simpson 3/8 rule is:

$$I = \frac{3h}{8}[f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)], h = \frac{b-a}{3} = \frac{x_3 - x_0}{3}$$

In addition to the methods mentioned above, there are many other aspects, such as Linear Equation, or Curve Fitting. Each algorithm in numerical methods is used for a specific purpose because its applicability is undeniable. Because it covers many fields, as well as quite complex formulas, along with requiring the use of a loop method to solve. Numerical methods is an extremely difficult subject to learn well.

2.1.4 Probability, Statistics and Random Process

"Probability" is a branch of mathematics that examines the possibility that an event will occur. It entails calculating the likelihood of an event occurring or not occurring based on available information [38]. Several applications in the field of IT make extensive use of probability theory can be followed by variating the appliances of Artificial Intelligence (AI), such as natural language processing, image processing, and robotics. While "Statics" has been applied in the Data Science area to analyze and interpret data, its main function is to draw a conclusion from collected data and predict what's next. Besides, it is an efficient tool for helping the bots figure out large data with random noise
Some outstanding concepts of Probability, Statistics, and Random process that we can apply to solve technology-related problems:

- Hypothesis Testing: is a statistical method used to test a hypothesis or claim about a population parameter using sample data. The formula can be used to implement is $Z = \frac{x-u}{(\frac{SD}{\sqrt{n}})}$
- Bayes' Theorem: is used to calculate the conditional probability of an event based on prior knowledge of related events. There are formulas for implementation:

$$P(Y|X) = \frac{P(X,Y)}{P(X)}$$

$$P(Y|X) = \frac{P(X|Y)P(Y)}{\sigma_Y P(X,Y)}$$

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X|Y)P(Y) + P(X|\ Y)P(\ Y)}$$

- Counting and conditional probability are important concepts in probability theory and statistics. The likelihood of an event happening is examined in conditional probability based on the likelihood of an earlier event happening. The first event determines the second event.

2.2 Applied Technologies

2.2.1 Mobile app technologies

Understanding the importance and complexity of the above-mentioned forms of math, creating a system to support students in learning math is necessary. As mentioned, this thesis introduces AiMA, an application developed based on a mobile platform. To deploy applications on this platform, several typical technologies are applied. The technologies used must meet the following criteria:

- Easy to deploy: it reduces application deployment time, therefore reducing time to release and providing quick solutions to users. Furthermore, this factor often comes with effective system management and update tools, making it easy for developers and administrators to develop continuously and maintain applications.

- Low development costs: making the cost of the application more affordable for users. Increase access to more customers, especially pupils and students, because they have low-income sources.

- Suitable for the nature of the project: the user's main task with this application is calculation, so choosing an application to optimize the calculation system is necessary. This helps reduce calculation time and response time to users, creating a better user experience.

- Feasible in deployment: the purpose of this application is to reach the end user, so deployment is the issue that needs the most attention. In addition, cross-platform and multi-operating system deployment is also a goal set for the entire application.

Our system applies MVC (Model-View-Controller) architecture. This is a powerful and popular software architecture model that enhances the organization and ease of maintenance of applications. This model separates the application into three main components: Model, View, and Controller.

- Model (data): this is the part that has the function of storing all the data of the application and is the bridge between the two components below, the View and the Controller. A model is the data used by the program. This can be a database, a normal XML file, or a simple object.

- View (interface): this is the user interface. View is the means of displaying objects in an application. For example, displaying a window, button, or text in another window. It includes anything visible to the user.

- Controller (controller): this is the part responsible for handling user requests sent through the View. A controller includes both Model and View. It takes input and performs updates accordingly.

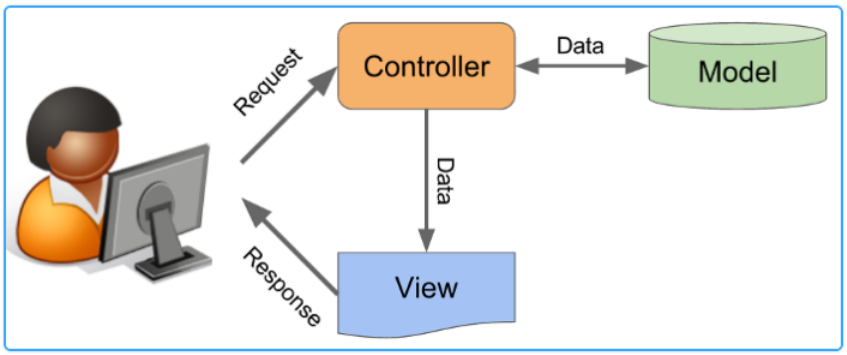


Figure 2.2: MVC architecture

Users will interact with the system through the interface side. In the math application, their action will be to enter the data of the problem, then using the network protocol, the data will be sent to the computer side. topic to solve the problem and after the calculation is complete, the results will be returned to the user. Moreover, a forum functionality is applied in this system for another purpose, which is about connecting all users for knowledge exchange. In addition, it is possible to consider the corresponding relationships between components of the MVC (Model-View-Controller) architecture and other components in a common application that uses a client-server architecture and has a database [39].

2.2.2 Artificial intelligence model technologies

Hybrid Vision Transformer (ViT) architecture

Hybrid-ViT, short for "Hybrid Vision Transformer," is a paradigm in the field of machine learning and artificial intelligence. This model combines convolutional neural network (CNN) and Vision Transformer (ViT) components for image processing and analysis. In the Hybrid-ViT model, the first CNN layers are used to extract features from images. These features are then fed into a ViT network, where they are processed through a self-attention mechanism to gain a deeper understanding of the image. By combining both methods, Hybrid-ViT aims to take advantage of CNN in local feature extraction. This improves the model's performance in many tasks, including optical character recognition. There are reasons why this model is more optimal for the OCR problem than the traditional ViT model, specifically:

- Although ViT is very effective in modeling complex and comprehensive relationships in image data, it is not always effective in extracting local features - an important factor in symbol recognition. CNN, on the other hand, is very powerful at extracting local features such as edges and corners that are necessary for character recognition. This combination in Hybrid-ViT ensures that the model can understand both the general and local details. Hence, the accuracy of the model will achieve a significant improvement.

- OCR often has to deal with images that are noisy or distorted. CNN in Hybrid-ViT helps the model resist noise and recognize characters in image variation conditions, while ViT contributes to analyzing the relationship between characters and the context around them.

Therefore, in this section, each part of the model will be analyzed.

Backbone

ResNet, short for "Residual Network", is a type of deep neural network architecture introduced in the paper "Deep Residual Learning for Image Recognition" in 2015. It is one of the network architectures The most popular and effective neuron for computer vision-related work. The highlight of ResNet is its residual block structure, where each block learns the "residual" of the function, i.e. the difference between the input and the desired output. ResNet can train networks with depths of up to hundreds or even thousands of layers without encountering the problem of vanishing gradients, a major challenge with previous deep neural networks. For the fine-tuned model, named pix2tex, it uses a custom ResNet layer, named ResNetV2. This backbone inherits the above strengths of traditional ResNet, as well as tweaks many parameters to make it more suitable for the task of identifying mathematical equations and formulas. This will be stated more clearly in the methods section.

Vision Transformer (ViT)

Transformer models are often used for NLP (Natural Language Processing) problems because of the self-attention mechanism, which helps the model focus on important parts of the input, which helps the model process handle complex inputs. Additionally, Transformer can be easily scaled up to the number of layers and the size of each layer, which makes it suitable for highly complex NLP tasks and large data. In addition to NLP, Transformer can also be used for computer vision problems. However, this model is still not optimized because there are many limitations. Therefore, most programmers will still prioritize convolutional architecture for recognition problems, such as image classification, object segmentation, etc.
To find a new, more effective direction for image recognition models, Google's research team introduced the Vision Transformer (ViT) architecture, which is a great computer vision version of the Vision Transformer (ViT) model. Transformer image. This model was introduced in the article "AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE" ¡ref¿. ViT has proven its performance with many outstanding results not only on classification problems but also on optical character recognition (OCR) problems.
First, about ViT's initial architecture for the image recognition model, the entire process is shown specifically in the image below.

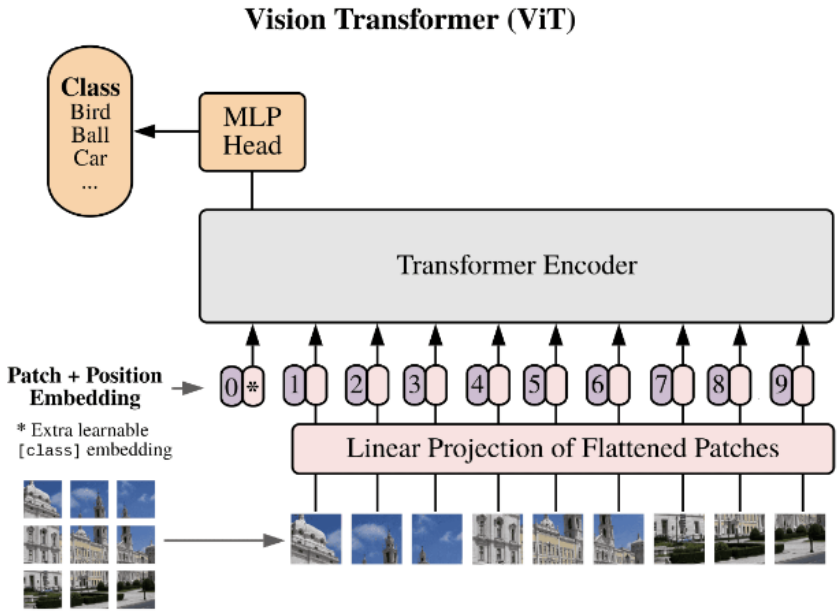
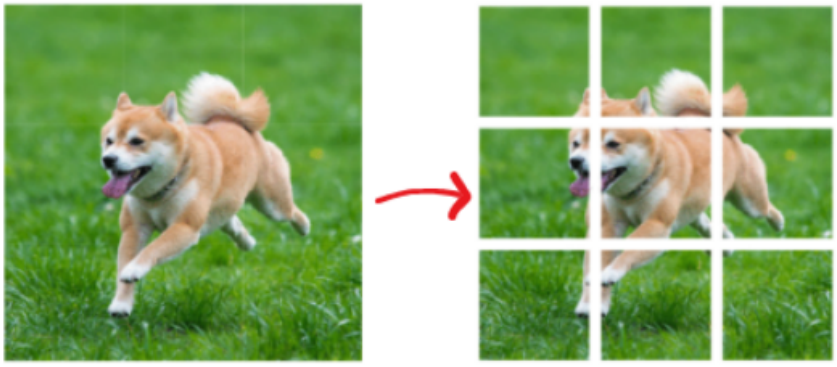


Figure 2.3: Vision Transformer architecture

The image above shows how the input image will be processed using ViT architecture. The process includes 6 steps, specifically:

1. Divide the image into same-size patches

   Recalling the Transformer architecture, its input is a 1D embedding token sequence. Therefore, for the Vision Transformer model, to process images with 2D dimensions, the ViT model will divide the input image into parts of equal size (usually 16x16) called patches. The number of patches can be calculated by the following formula: $N = \frac{HW}{P^2}$, with H, and W as the dimension of the input image. So the initial dimension of the image can be performed by $xR^(HWC)$. After the division, it will be $xR^(N(P^2C))$, where H, W, C, N, and P are height, width, channel of color, number of patches, and divided dimension respectively. The below image will fully illustrate this idea:



$$P = 16 \quad H = 48 \quad W = 48 \qquad N = \frac{HW}{P^2} = \frac{48 \cdot 48}{16^2} = 9$$

Figure 2.4: Patchifying the input image

There is a definition is embedding size D, which is the number of dimensions of the vector representing each patch after being embedded into the model. Because VIT still uses the constant embedding size D in the model, we need a linear transformation that turns the packets to size D. This transformation is also called patch embedding.

Next step, convert these patches into vector form by flattening these patches. A fully connected layer has its responsibility to flatten the patch into a vector. For details,

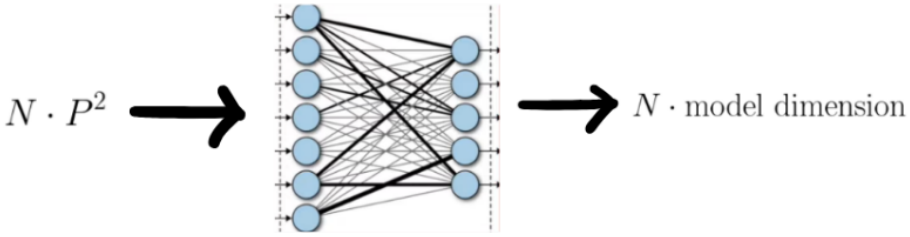


Figure 2.5: Vector flattening

A simple fully connected network is applied to all patches, so the output of this step is a 1D sequence of size calculated by several patches time with the model dimension. Then, linear projection will be applied. Linear Projection is essentially a Dense class with the input being the flattened vector of patches, the output will be the embedding vector corresponding to each patch. The formula here will be: $z_n = Wx_n + b$, where x_n is the flattened vector of the n^{th} patch, z_n is the corresponding output of x_n through Linear Projection, and W is the embedding matrix. The figure below is the full description of this process:

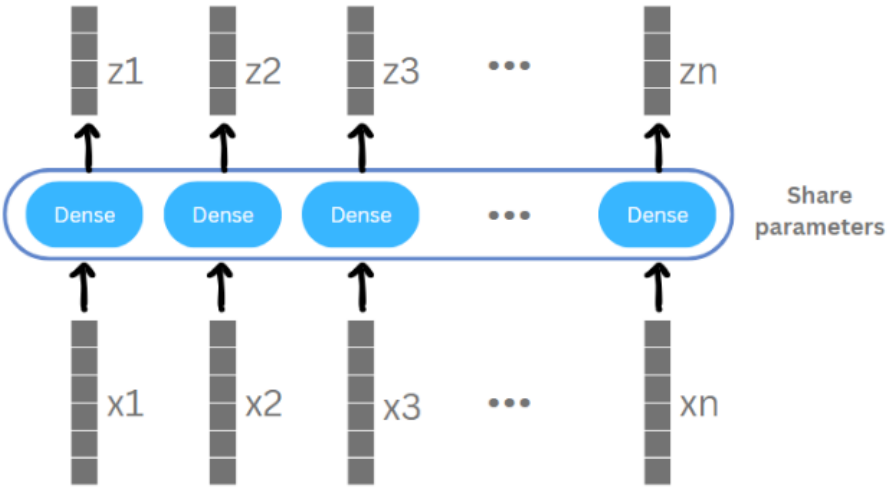


Figure 2.6: Linear Projection description

2. Position Embedding

   In the Vision Transformer (ViT) model, the use of Position Embedding is important because the model is unlike traditional models using CNN (Convolutional Neural Network), where the spatial location of features is preserved through pooling and convolution layers. If just embed the patches and put them into the model, it will face the above problem.

Figure 2.7: Positional problem in ViT

   If just embedding the patches and putting them into the model, there will be no difference between the two images above. Therefore more information about the location of each patch is necessary. Positional embedding is a way for a model to learn about the location of information in a sequence or multidimensional space. When working with data that does not have absolute order or position information, such as in images, positional embedding helps it understand and retain information about the positions of elements.

3. Transfer Encoder

   The transformer encoder receives input that consists of the combination of patch embedding and position embedding information, which includes Multihead Attention layers, MLP block, and Layer norm. The transformer encoder can be performed by the below figure:

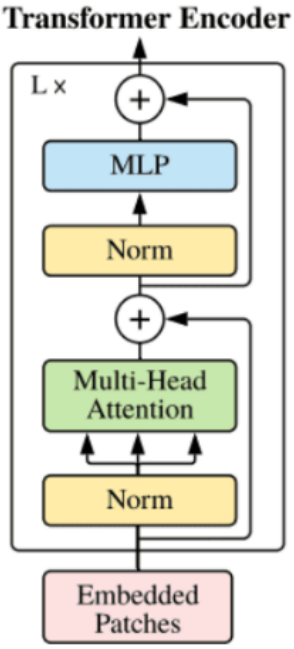


Figure 2.8: Transformer Encoder architecture

- Norm ¡ref¿: norm here is layer norm (LN), allowing the model to adapt to variations in training sets. In addition, this is also a method to improve training speed with diverse neural network models. In the ViT model, the norm layer is used to normalize the input data and the data after performing self-attention steps in the self-attention layer. This helps maintain mean and variance across layers, making the training process more stable. They can be calculated by

$$\mu_i = \frac{1}{m}\sum_{j=1}^{m} x_{ij}$$

$$\sigma_i^2 = \frac{1}{m}\sum_{j=1}^{m} (x_{ij} - \mu_i)^2$$

where m denotes the number of hidden units in a layer. The layer normalization method can be displayed below:

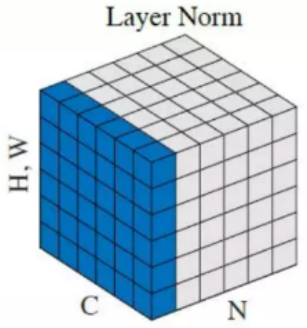


Figure 2.9: Layer Norm

where N is the batch axis, C is the channel, and H, and W as height and width, respectively. The blue squares are normalized by the same variance and mean. They are of equal size, indicating that the features have been stabilized.

The self-attention layer is the main part of generating a block in the Transformer Encoder. The input of the self-attention layer is the vector sequence obtained from the previous step. Suppose it is the sequence $X = [x_1, x_2, x_3, ..., x_n]$. It will be passed through the self-attention layer, specifically:

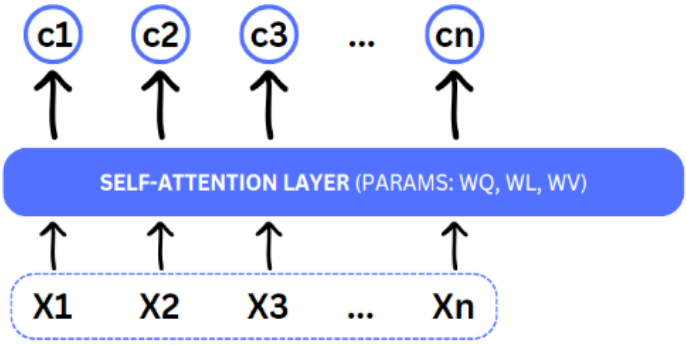


Figure 2.10: Self-attention layer

The idea of this class is to create a context vector C containing the most important information of the input string $C = [c_1, c_2, c_3, ..., cn]$. In the self-attention mechanism, three sets of parameters W_Q, W_K, W_V are used to transform the input vector into three query vectors, key vectors, and value vectors.

- Query vectors: This query vector will be used to measure the similarity between the current element and other elements in the image space.
- Key vectors: This context vector will also be used during computation to measure similarity and generate attention weights.
- Value vectors: This value vector will be used to create a weighted sum of the values corresponding to each query vector.

Below is how the above 3 vectors create the C context vector:

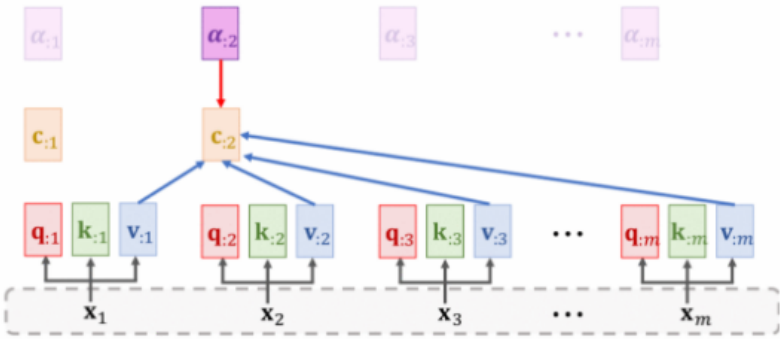


Figure 2.11: Generating the context vector C

Behind this layer, there is a way to calculate the above values:

- With each value of x_i of the input string X, calculating q_i, k_i, v_i by the formulas: $q_i = W_{Qx_i,k_i} = W_{Kx_i,v_i} = W_{Vx_i}$

- Calculate the alignment score with corresponding x_i follow the formula: $\alpha_i = Softmax(K^T q_i)$
- Finally, calculate the context vector C by: $c_i = \alpha_{11} v_1 + \alpha_{21} v_2 + ... + \alpha_{m1} v_m$

- MLP: MLP stands for (Multi-Layer Perceptrons). It is a neural network consisting of two layers that utilize the GELU activation function at the final layer. The last Multi-Layer Perceptron (MLP) block, also known as the MLP head, serves as the output of the transformer. If the result is used for Image Classification, applying softmax can yield classification labels. It is shown by the figure below:

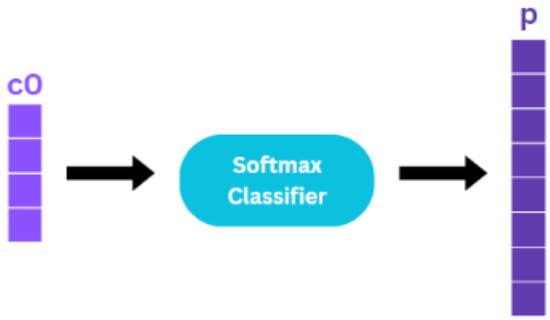


Figure 2.12: Multi-Layer Perceptrons for classifying images

This block receives as input the context vector c returned from the self-attention layer and gives the final result the probability corresponding to the classes.

However, this is just how the Vision Transformer model works for image classification and recognition problems. ViT needs to be fine-tuned to suit the characteristics of each different problem. Specifically, for the optical character recognition problem, ViT will be used as a completely new version of it. That is Hybrid-ViT, a hybrid architecture, details of which will be presented in the next section.

2.3 Related works

Because this topic is not too new, there have been quite a few projects implemented. However, not every project can bring the best results. Through the testing and evaluation process, these models still have many problems that need improvement.
The problem that most models encounter is similar to Khiem's model. Khiem T. Do and his team use a convolutional neural network for OCR tasks. The strength of this model is its ability to recognize many typefaces, such as handwriting and print. The main dataset this model uses comes from the MNIST dataset, which is a familiar dataset in the field of machine learning and computer vision. This dataset is often used to evaluate and compare the performance of machine learning models in the problem of handwritten digit recognition. Although the data set used is quite large, with many classes on different and diverse mathematical characters, however, this model can only work well on numbers and basic mathematical characters. such as plus, minus, multiplication and division. Additionally, this model can only handle simple background problems, such as pictures of black text on a white background. This can be explained by CNN not performing very well in this case because CNNs can be affected by noise or complex backgrounds in

the optical image. Besides, because it does not use the self-attention mechanism, CNN can only learn features such as edges, corners, or shapes of characters, but it does not understand the relationship between them. In other words, CNN has difficulty focusing on the important part of the image. This makes it impossible for the model to recognize complex equations such as those with square roots, exponents, or integrals and derivatives. This is especially inconvenient for the majority of users when they are mainly interested in more complex and difficult-to-solve problems.

In contrast, Lukas Bletcher's pix2tex model proves effective when it can recognize more mathematical symbols well and can recognize equations of higher complexity. However, currently, the model cannot recognize images of the user's handwriting. Therefore, this is the point that this research focuses on solving.

Lukas Bletcher's model called pix2tex is a model that can solve the above problems. This model uses ViT architecture combined with CNNs, so it can inherit the strengths, as well as overcome the limitations of a pure CNNs model mentioned above. The pix2tex model can be recognized from simple problems such as basic math operations, and first-order equations, but can also understand more complex equations, such as 4th-order, 5th-order equations, and 1st-order equations. Programs containing radicals, or problems related to calculus. However, by using the ViT model as the main development architecture, data diversity is essential, but this model still does not really work well. According to Lukas Bletcher, this model was created to help users copy problems, equations, and mathematical formulas from images as quickly and effectively as possible. So, the main data of the model comes from the im2latex-100k dataset, which includes nearly 100,000 images of printed equations. This helps the model to only be able to recognize images of equations written from a computer well, but for handwritten equations, it is completely unrecognizable. In addition, a weakness of the model is that it is not good at recognizing images with complex backgrounds (non-white backgrounds), equations that are not neatly written, as well as equation images with low pixels. According to personal assessment, this is a potential model for further development, so the main purpose of this thesis is to apply fine-tuning on this model, to overcome the above weaknesses.

Besides, CoMER's model has shown remarkable performance when it can recognize complex equations, including handwriting, and this is also the strength that this research wants to achieve. Nevertheless, this model can currently only recognize black and white images, and input images with complex details, or colors, will be difficult problems for detecting.

A model by Anh Duy Le and his colleagues is also a model that uses Hybrid-ViT architecture that combines ViT and CNNs. Like Lukas Bletcher's model, the model is good at recognizing complex equations, however, according to the report, there is no information for predicting handwritten math problems, but analyzing the dataset used for the application mentioned in the article, the data set is mainly printing formulas. Therefore, it is understandable that this model's ability to predict handwritten characters will be limited. In addition, according to the Future Work section of the article, this model is only being implemented as a research article and is not yet capable of being applied in practice. This is also a factor that Anh Duy Le and his team will consider to improve in the future.

In terms of a mathematical application, Photomath from Microblink company provided a mathematics solver system, this application is only supported on mobile platform. This is one of the most popular applications in this branch, therefore in 2017, the Tech Edvocate named Photomath among its top 20 teaching and learning apps, and in 2021, the legally constituted business Photomath LLC reported that it had raised 23 million in Series B

funding, with Menlo Ventures serving as the lead investor. Photomath is a powerful tool that can solve problems in the only field of calculus. It can quickly find solutions to issues involving integrals, derivatives, etc. However, Photomath has difficulties in solving long and complex expressions. However, a user-friendly interface makes Photomath available to more people. Photomath has a team of R&D (Research and Development) to research mathematics methodologies all around the world to supply the correct solution and ensure all the shown steps are expert-verified. With sensitive AI techniques, Photomath uses the camera on a user's mobile device to scan and recognize mathematics symbols for solving it. In the early years, Photomath could not recognize the handwriting problem, but from 2016 until now, the app began recognizing handwriting in addition to printed text. The fact that Photomath can demonstrate how to solve each problem totally for free is a significant benefit, but the explanations are limited. Due of the student feedback collected, Photomath can be used in the actual classroom for teaching purposes. In addition, students stated that Photomath provides a visually appealing computational assistance tool with straightforward interaction. Likewise, Israeli businessmen Michal Avny, Adam Arnon, and Lev Alyshayev launched to the market a mathematics mobile application called Symbolab in 2011. This is a multi-platform application, which can support users on both mobile and web platforms. However, for the best experience, as well as being able to use all features of Symbolab, users are encouraged to use the application on mobile devices when the AI scanner is only available on this platform, users can use the camera on their mobile device for detecting the expected mathematics problems. A constraint is that it cannot eliminate redundant characters. Like Photomath, Symbolab is able to solve problems in calculus, however, this application has a variety of math topics including geometry, trigonometry, and linear algebra. Besides, Symbolab's step-by-step display system is rated as effective, therefore, it has been suggested to be employed as a teaching tool in various research, including the research paper of Nguyen Viet Duong et. al. Mathway is a different, equally well-liked application. Mathway, which builds on earlier apps, has a far wider range of disciplines under its belt. In addition to math, Mathway can support physics and chemical formulas. The capacity to answer trigonometrical issues is the most notable, and Mathway has been successful in this area since it offers many formulas in addition to intuitive use. Like Symbolab, Mathway is a multi-platform application, besides, possessing a powerful AI model that can detect handwriting with very high accuracy.

To summarize, our proposed model was created to inherit the advantages and overcome the disadvantages of previous models. This model will be able to recognize a variety of image types, such as images with many complex details, and images that are skewed, compressed, or blurred. In addition, to be accessible to students of multiple grades, the model can recognize not only equations but also complex specialized formulas. Realizing that, to make the model more valuable, integrating the model into a math application, to support students and students in learning math is a wise decision. And finally, recognizing handwriting is also the goal that the research team is aiming for. Thus, this is intended to be a complete model that can be integrated into many practical applications, with different purposes.

Chapter 3

METHODOLOGY

3.1 Mobile application development

Summary of the proposed system, as mentioned, AiMA is implemented mainly based on the mobile platform. With this platform, the application can be easily accessible because of its convenience and ease of access anywhere and everywhere. The mobile system is based on the MVC architecture for development, to distribute tasks into two main sides to avoid unilateral processing, which can cause system overload and optimize the application's maintainability. The client side will play the role of displaying the interface, as well as receiving requests from users. The server side will process those requests, and respond with the user's desired data.
Because this is a complete software system, it will be necessary to apply software development models, which help enhance work organization and effective project control. The clear division into stages in the model helps the project to be organized systematically, from defining requirements to implementation and maintenance. This mobile system applies a classic software development model, called Waterfall. This model describes the software development process as a series of linear steps, following a single flow from start to finish. The waterfall development process model consists of several distinct phases, including requirements gathering and analysis, design, implementation, testing, deployment, and maintenance. Each phase must be completed before moving on to the next one. The figure below is the general observation about this model:

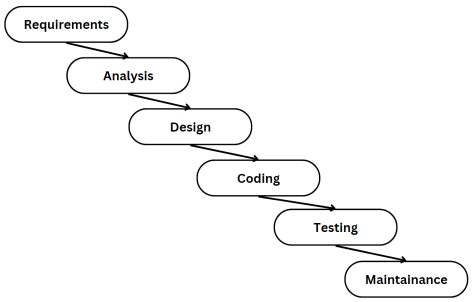


Figure 3.1: Waterfall model

3.1.1 Requirements and Analysis

In the Requirements and analysis stages, the project requirements were collected, analyzed, and documented. For details, to start building any software, gathering requirements

from users is always paramount. The "Requirements and Analysis" phase in the software development process plays an important role by focusing on identifying and analyzing system requirements. During this phase, the development team focuses on gathering detailed information from customers and considering their needs and desires. Hence, conducting a survey can collect the potential customers' needs. In this research, a small-scale survey was undertaken to learn what the Information Technology and Logistics majors at the International University - Vietnam National University Ho Chi Minh City, Vietnam, wanted from the use of mathematics. The figure below indicates students' interest in areas of need. In this survey, the students from these majors are asked one question. The survey results indicated that the students from both groups of majors already knew and often used the mathematics learning support applications. From the published mobile applications, which can be installed from CH Play or App Store, such as Mathway, Symbolab, etc. The students responded that the main limitation of many current apps is still the lack of support for specialized mathematics fields, therefore; they want an application to meet this problem. The following is a result from this survey, which shows the most concerned specialized fields:

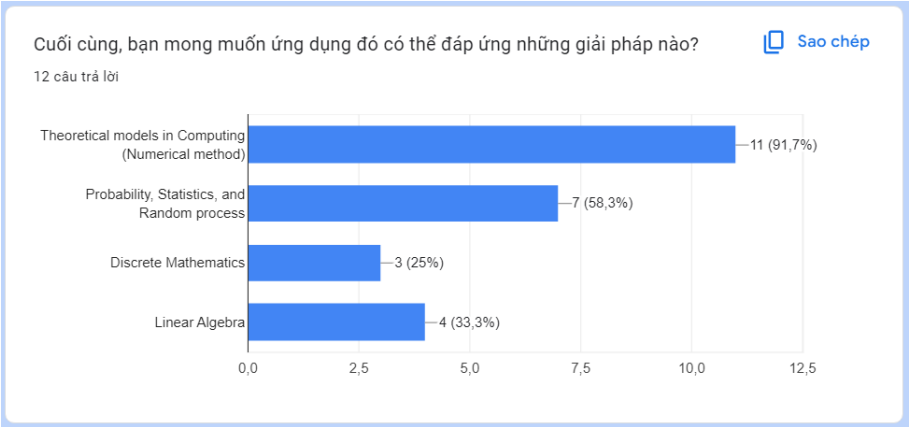


Figure 3.2: The bar chart indicates students' interest in areas of need

Based on the data collected from users, we were able to determine the direction and scope of the system. The above survey reflects that, besides the popular fields of mathematics, students are very interested in and wish to receive support in other advanced mathematics fields such as numerical methods and probability.
In the field of Numerical Methods, interested students are assisted in applying numerical methods to solve complex problems. The application can provide arithmetic calculation tools, algorithms, and illustrative examples so that they can clearly understand and apply this knowledge to real-life situations. This is understandable because, for students, especially engineering students, the Numerical methods subject is considered an important subject because of its application in system optimization.
In the field of Statistical Probability, students seek assistance in understanding and applying the concepts of probability and statistics to complex tasks. Applications can provide statistical tools, for example, their application to fields of research, business, and especially technical fields such as industrial engineering and computer science.
Therefore, the system will focus on developing algorithms related to the above two areas, in order to meet and reach as many users as possible.

3.1.2 Design and Implementation

Design and Implementation are two important stages in the software development process, they shape the structure of the system and convert ideas into actual source code.

Design system - Sequence diagrams

Regarding system design, it is necessary to create diagrams to represent and understand the interactions between objects in the system in chronological order. We used a sequence diagram, which is a powerful modeling tool in UML. It allows us to identify concurrent relationships between objects, and clearly understand the roles and responsibilities of each component in the process of performing a certain task or function.
As described, AiMA is a computational support system, that integrates AI technology along with providing a forum platform to help link users together. Therefore, the sequence diagram of the system can be designed by dividing it into two large parts, including authentication and the rest of the system. The below figures indicate the way this system works:

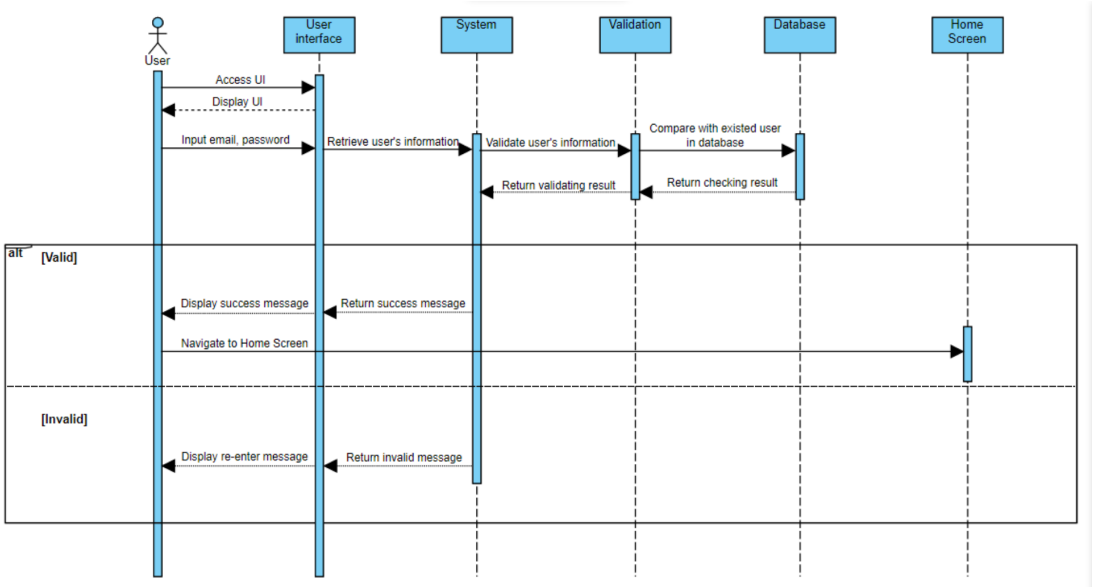


Figure 3.3: The sequence diagram of login functionality

The above sequence diagram describes the user sign-in process in the AiMA application. Below is a detailed description of the steps shown:

- User Interface: This is where the user interface is displayed to users, from which the system can receive their requests.
- User: The process starts with the user accessing the user interface, the user is required to enter their information about the account's email and password.
- System: the server of the system will receive this input from the client and start validating this data.
- Validate: The system sends this information to the authentication checker to verify the user's information.

- Database: The authentication check compares user information with existing email in the database. Besides, check the format of the input (the input can not be empty, and the password's length is at least 6 characters), and then send the authentication results to the system.
- System → User Interface: The system sends status messages based on the test results to the user interface.
- User Interface → User:
  - If the information is valid, the user interface displays a success message and then leads the user to the Home Screen.
  - If the information is invalid (Invalid), the user interface requires the user to re-enter the information by displaying an error message.

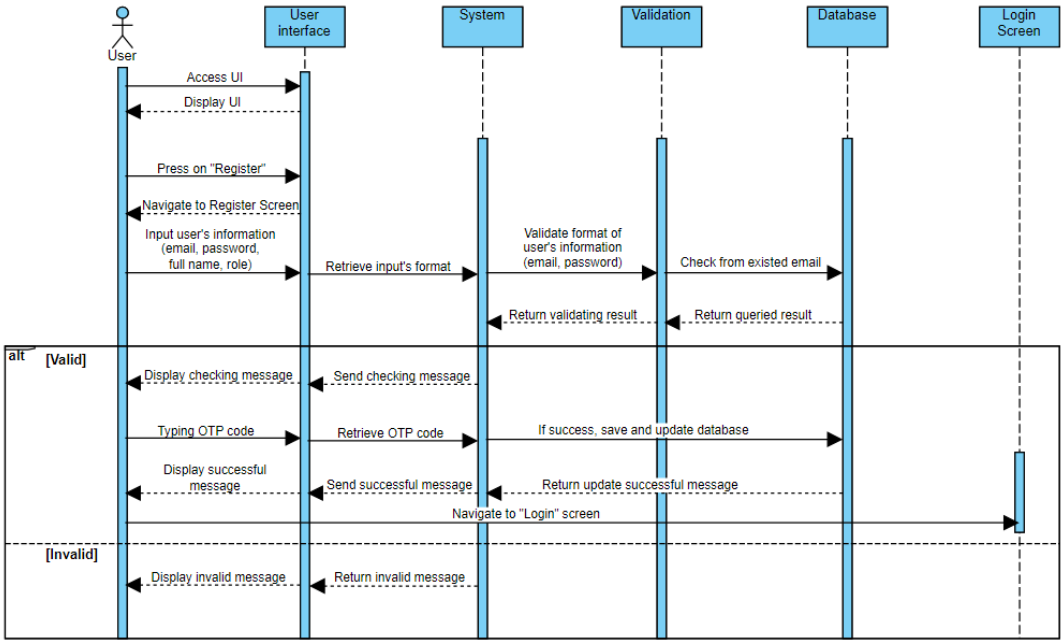


Figure 3.4: The sequence diagram of register functionality

The above sequence diagram describes the user sign-up/register process in the AiMA application. Below is a detailed description of the steps shown:

- User Interface: This is where the user interface is displayed to users when they press the "Register" button, from which the system can receive their requests.
- User: The user is navigated to the registration screen, where they enter personal information such as email, password, full name, and role.
- User Interface → System: The user interface sends this information to the system to check the input based on the correct format.
- System → Validation: The system sends information to the validation department to check the user information.

- Database: The validation department sends a request to the database to check whether the email exists or not. Besides, check the format of the input (the input can not be empty, and the password's length is at least 6 characters), and then send the authentication results to the system.
- Database → Validation: The database returns checking results.
- Validation → System: The validation department sends checking results back to the system.
- System → User Interface: The system sends checking messages to the user interface based on the status of the checking result.
- Checking result is valid:
  - System → Database: The system will save and update user information in the database.
  - User Interface → User: The user interface displays a message informing the user that the information is being checked. The user then enters the OTP code for authentication and the user interface displays a success message
- Checking result is invalid:
  - Do not save and update the database.
  - The user interface displays an error message to the user.
- User → Login Screen: The user will be navigated to the login screen to start the sign-in session.

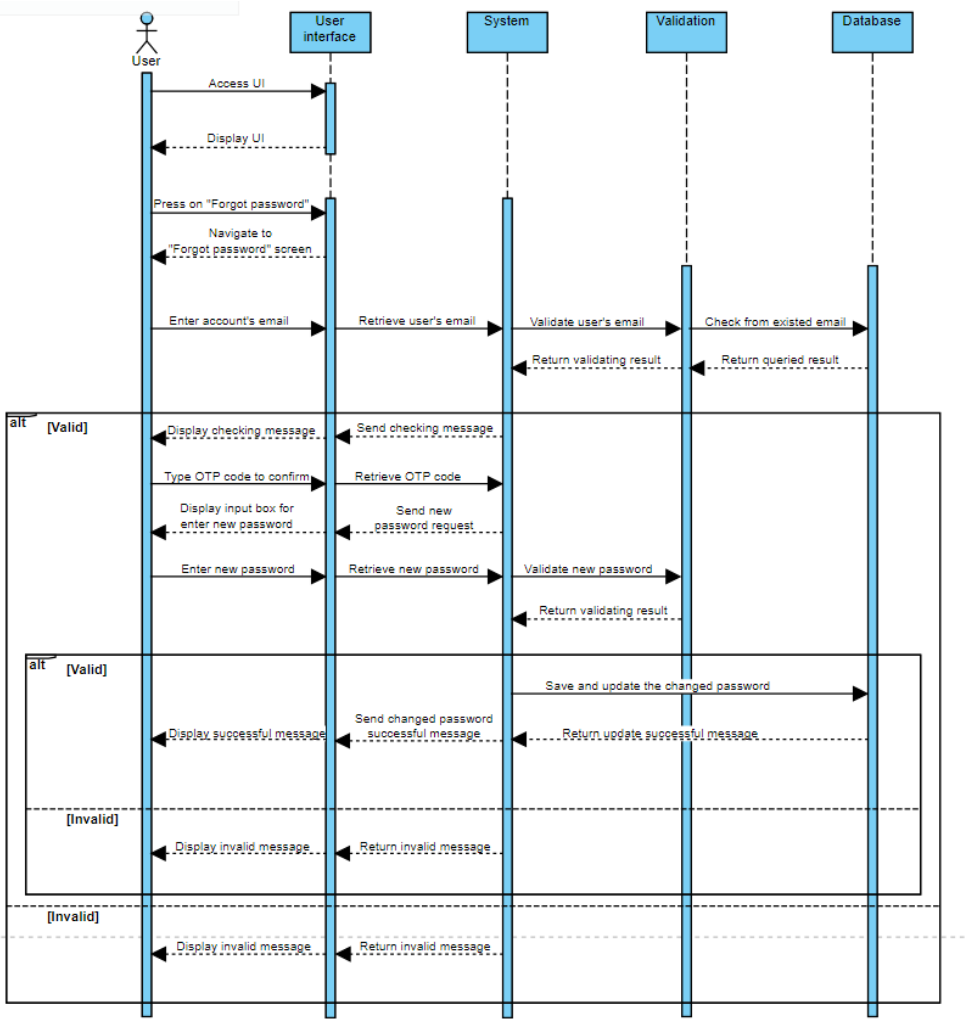


Figure 3.5: The sequence diagram of forgot-password functionality

The above sequence diagram describes the process when the user presses the "Forgot Password" button in the AiMA application. Below is a detailed description of the steps shown:

- User Interface: This is where the user interface is displayed to users when they press on the "Forgot Password" button, from which the system can receive their requests.
- User: The user is navigated to the "Forgot Password" screen, where they enter their account's email to reset the password.
- User Interface → System: The user interface sends the user's email to the system.
- System → Validation: The system sends an email to the validation section for checking.
- Validation → Database: The validation department sends a request to the database to check if the email exists.
- Database → Validation: The database returns validating results.
- Validation → System: The validation department sends validating results back to the system.

- Email validation result is valid:
  - System → User Interface: The system sends successful messages to the user interface.
  - User Interface → User: The user interface displays a message informing the user that information is being checked.
  - User → User Interface: The user then enters the OTP code and then enters the new password.
  - System → Validation: The system sends the new password to the validation section for checking.
  - Validation → System: The validation department returns the validating result.
  - New password validating result is valid:
    * System → Database: If the new password is valid, the system requests the database to save and update the password.
    * Database → System: The database returns a successful update message to the system.
    * System → User Interface: The system notifies the user interface about the successful update.
    * User Interface → User: The user interface displays a success message to the user.
  - New password validating result is invalid:
    * System → User Interface: The system sends error messages to the user interface.
    * User Interface → User: The user interface displays an error message to the user.
- Email validation result is invalid:
  - System → User Interface: The system sends error messages to the user interface.
  - User Interface → User: The user interface displays an error message to the user.

Use Case Diagram is a type of UML (Unified Modeling Language) diagram used to describe a software system from the user's perspective. It helps define interactions between users (also called "actors") and the system to achieve a specific goal. Through diagrams, developers, end users, and stakeholders can easily discuss and agree on required business functions and processes. In addition, diagrams assist in building the structure and design of the system's user interface, ensuring that it accurately reflects and supports the user's workflow. The diagrams, that are shown below, are often used in the analysis stage to determine the actor's ability to perform the function through specific use cases. Users may interact with the system by using computing features as well as CRUD-related activities (create, read, update, and delete) when using the Forum feature. Similarly, the system can take other actions like altering support algorithms or controlling user posts and information. Administrators may be able to better manage their apps with this explicit allocation.

Design system - Use case diagrams

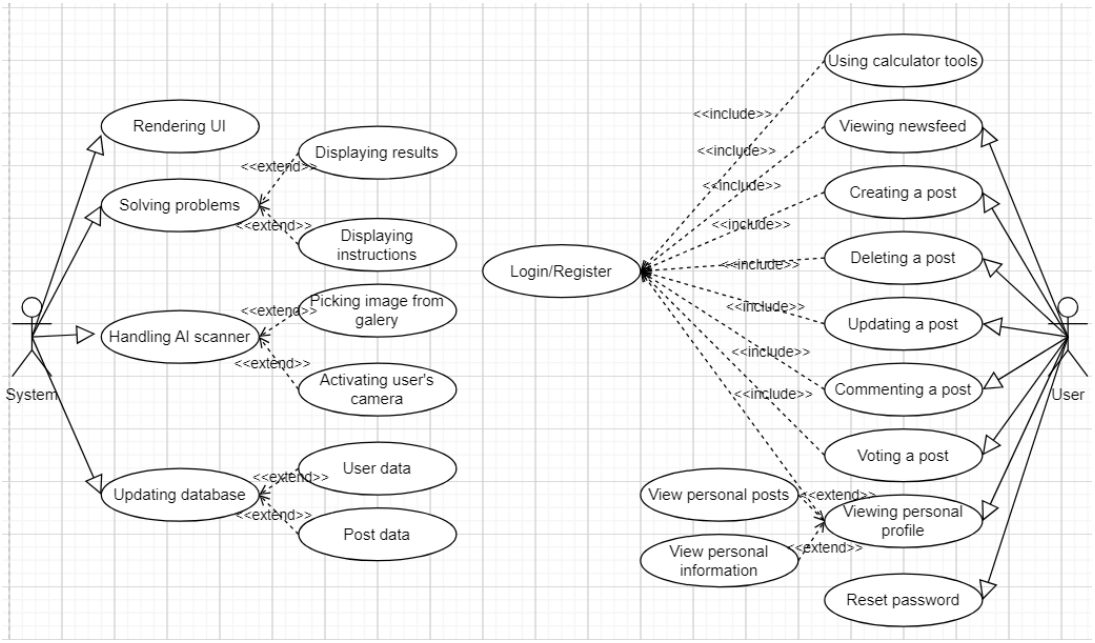


Figure 3.6: Use case diagram

The Use Case diagram above provides an overview of the functions that the AiMA system provides and how users interact with these functions. The diagram includes two main actors: "User" and "System". Here are some highlights from the chart:

- System: Represents the underlying system, with functions such as displaying the user interface, solving and calculating user problems, and displaying corresponding results and implementation steps. Regarding the mathematical equation recognition feature, camera processing and selecting images from the library are the user's choices. Along with that come methods for updating the database, specifically for user and post data.
- User: Represents the end user of the application, and has many interactions with the system, including logging in/registering, viewing message boards, creating posts, deleting posts, updating posts, commenting, voting for posts, viewing individual posts, viewing personal information and reset password.
- Login/Register: This is the main contact point between the user and the system. To perform certain features, specifically database-related features, users need to go through this process.

As can be seen, the typical features that users can interact with the system include creating posts, deleting posts, updating posts, commenting on their own or other users' posts, and upvoting these comments. These features will be made more clear in the system design section for the Forum feature.

Design system - Database design

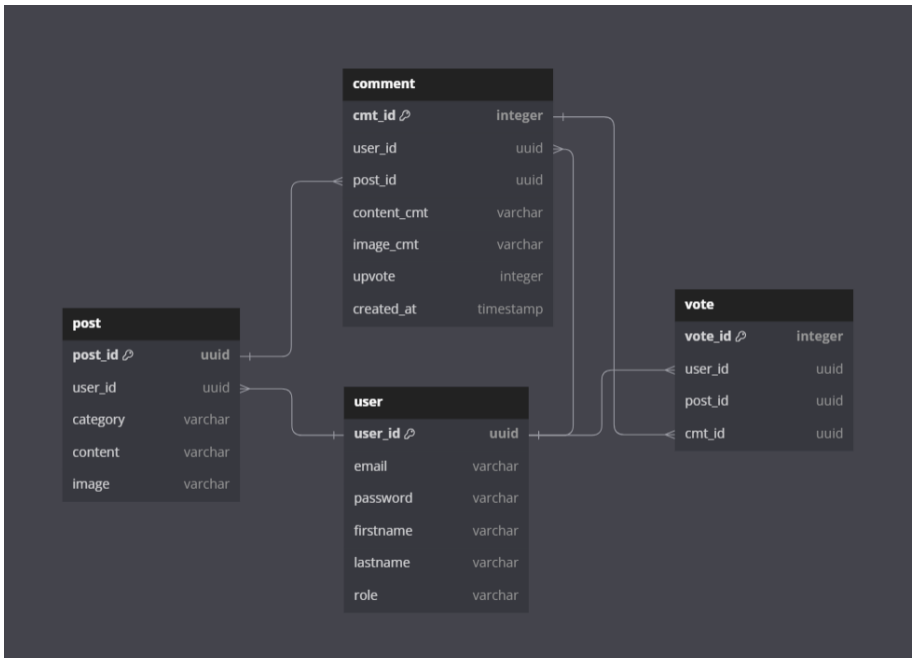


Figure 3.7: Database Schema

Designing a database schema is essential because it helps manage data in an organized and systematic way but also plays an important role in improving system performance. The image above is a database diagram of the project's forum features, including 4 main entities, namely user, post, comment, and vote. It works as follows.

- For each user, can own many personal posts, each post will carry its ID value (post ID), along with the ID of the poster (user ID).
- For each post, it includes title, content, poster name, posting time, and especially comments from users. Each post has a unique ID, called post ID, each comment will also include the comment ID, post ID (ID of the post this comment is in) along the ID of the person who posted the comment (user ID). For each post, there can be many comments, including comments from the owner and others.
- Each comment will carry an ID (called comment ID), in addition, each comment will have a value called upvote, which represents the number of people who like that comment. Each vote will also have the ID of the person who clicked, the ID of the comment, and the ID of the post it is in.

Design system - Process diagrams

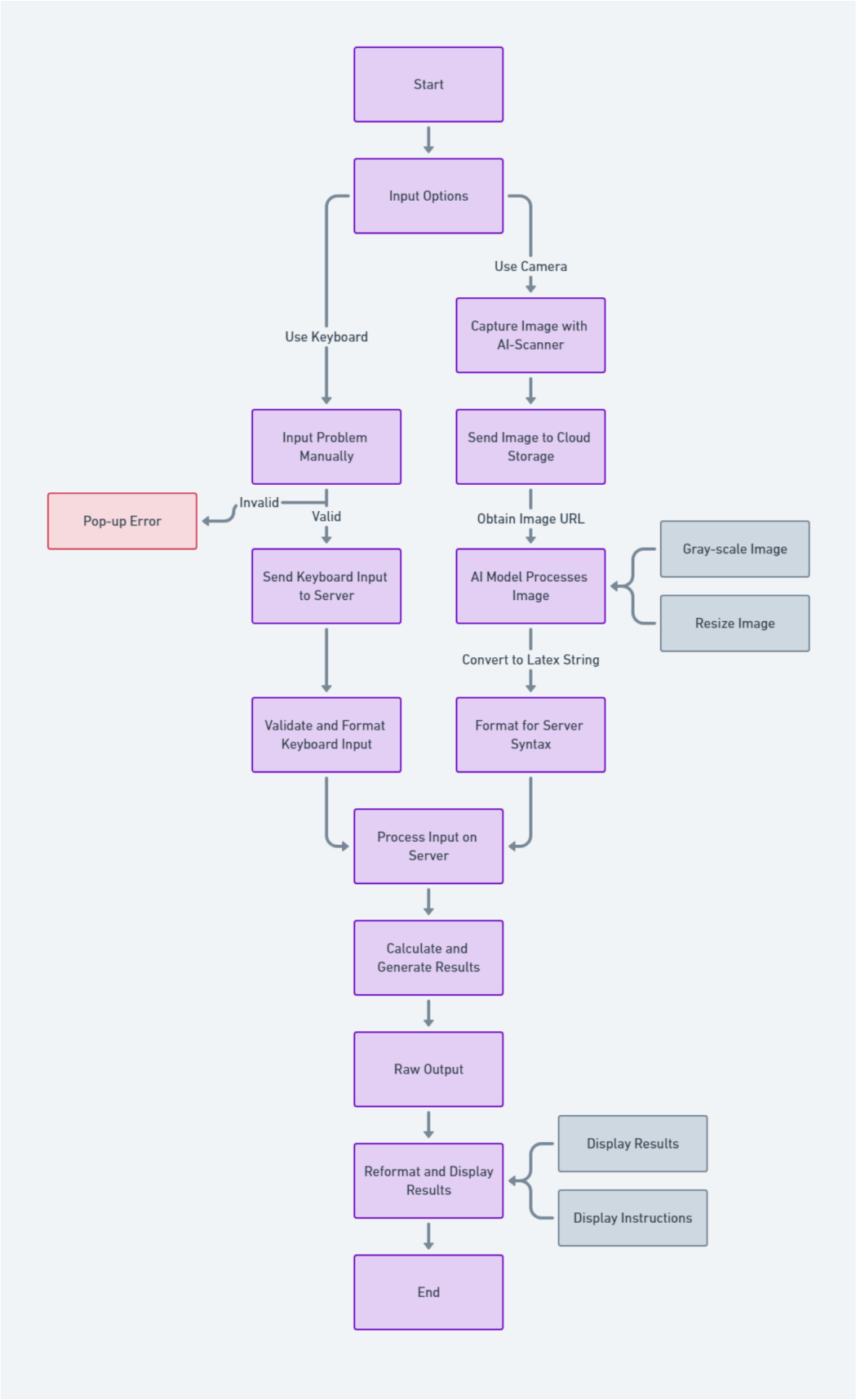


Figure 3.8: Process diagram

Within the AiMA system architecture, each client request must adhere to a certain protocol. The input, handler, and output components are usually the three primary components of a basic operation. Every component has a purpose. To be specific:

- Input process:
  - There are two choices available to the end-user:
    - Enter the mathematics equations using the system keyboard. Users may manually enter the required data into the client's provided input area.
    - Use the smartphone's camera for capturing the image of the equation: Users can get the result from the equation's image captured by their camera.
  - The following processes receive the input that is provided to the backend server.
- Handler process:
  - String input:
    - To make sure the input string is compatible and appropriate for computation, it must go through validation; if it is not, errors will appear.
    - The back-end server is in charge of determining the kind of string to route to the appropriate controller if the input is valid.
    - The format of the input will be adjusted to conform to the server-side programming language's syntax.
    - At this stage, the server is ready to run the calculator and provide the final output.
  - Image input:
    - To retrieve its URL, the picked or captured image will be saved in a cloud-based platform.
    - The URL can be read by the AI model. Error pop-up if the URL is not specified.
    - If the URL is valid, the AI model transforms it into a latex string.
    - The format of the incoming latex string has to be adjusted to conform to the server-side computer language's syntax.
    - Server can tackle the problem.
- Output process:
  - The mathematical transformation will modify the initial string that was received.
  - Direct the user to the "Solution" screen where the steps and formatted results will be shown.

Mobile Core Technologies

The implementation phase in the Waterfall model is the process of realizing the system based on the design determined in the previous phase. This is the process of moving from the design model to the actual source code, so determining the core technologies for the system is extremely necessary.

1. Client:

   On the front-end side, the system uses React Native as the main development tool. React Native is a mobile application development framework, helps developers build cross-platform mobile applications [12]. React Native quickly became an important tool for effective mobile application development. Choosing React Native for client-side development is a suitable choice because this framework has met the set criteria, specifically:

   - React Native uses JavaScript as the main programming language, this is an easy-to-use language because the syntax is transparent, easy to remember, as well as intuitive for programmers. Besides, React Native integrates various plugins and libraries, supporting users in many tasks, from building interfaces to methods of connecting to the server. In addition, React Native inherits almost absolutely the strengths of its predecessor, ReactJS, a JavaScript library for building user interfaces on the web app platform, this makes React Native more accessible to newcomers. Now, just by knowing how to use basic React, programmers can easily access application development on the app platform.
   - React Native provides extremely effective component reuse, which makes the project's code easier to see and manage [12]. In particular, a math application, requires many screens to be built, and the characteristics of the screens will usually not be too different, so reusing components is a good decision. In addition, React Native supports hooks, which help save cache and limit unnecessary function calls. From there, the application can operate more smoothly and effectively.
   - React Native provides multi-operating system (Android and iOS) deployment features, helping developers save time and effort. With a common source code, applications can be deployed simultaneously on both operating systems without requiring significant changes. This increases efficiency, reduces the possibility of errors, and maintains synchronization between application versions on different platforms, creating a consistent and quality user experience. In addition, because of the similarity in syntactic features, redeploying the system on another platform, specifically a web app, will be much easier, now we only need to focus on refining the interface to suit the website. The concepts of React Native include:
     - Component: Components are a fundamental concept of both React and React native. It is the division of applications into small components that creates their high reuse and scalability. In mobile applications, the header is the component that has a little bit of changes in the system. Therefore re-coding for it on different screens is an inconvenient job. Thanks to React Native, programmers only need to write code for the Header component once, and on other screens, just call it again. it's like a function to display.
     - Props: Prop is an abbreviation for "Properties". An important principle to keep in mind when utilizing props is that their values should never be changed. In other words, props are considered immutable data. Components are passed props from the parent component. Modifying the value of props in these components is prohibited; it can only be read. In React Native, data follows a unidirectional flow, specifically from parent components to child components.

- State: Props and State work in different ways. Props are the data that is sent to a Component, and state is the data that is stored inside it. This means the state can be fully changed, so we call it a mutable data type. The state is used to update the view's data and then bind the new data back to the view when there is a change.

2. Server:

   Python is a language that is no longer too new in software systems, especially software that requires robustness and stability on the server side in data processing [49]. A math application is no exception when its main task focuses on calculating and solving users' problems. Therefore, using Python is one of the top choices. Python can be evaluated by the following criteria:

   - Because Python is a modern programming language, designed for people to understand, Python is often known for its simple and easy-to-read syntax, helping developers understand the code and optimize the process. development process. This is important for building and maintaining code in the field of mathematics, where clarity and transparency are important.
   - As mentioned, the main computational task is paramount. Therefore, integrating support libraries is necessary, helping to save time and effort, as well as improve problem-solving ability. With Python, a language with strengths in statistics and data analysis, it is always fully equipped with libraries to support computing purposes such as Sympy, Numpy, Scipy or math, etc. These libraries provide the necessary tools for numerical computation, statistics, linear algebra, and equation solving, which are important elements in applied mathematics.
   - Python is also known as a popular language for artificial intelligence (AI) development because it provides many powerful libraries and frameworks, as well as a large community of support and rich resources. Therefore, using Python will help deploy AI models easier and faster.

   Some Python libraries used in this project, including:

   The main technology applied on the server side is the Flask server, which is a Python framework. With Flask-server, it can handle requests from the mobile application side. Built on the Python programming language, Flask is not only flexible and easy to use but also powerful with a large community support and a diverse plugin community. In addition, Flask-server also supports APIs for interacting with mobile applications, helping us easily expand features and integrate with other services.

   Moreover, forum-related functionalities, such as authentication, and CRUD actions can be handled on this side. For details.

   - Authentication:
     - Login: Users will enter information about their account, including email address and password, to access the application's forum feature. The system will query the user from the database.

```
result = QueryParamFunc('SELECT * FROM "user" WHERE email = %s', (email,))
```

     Listing 3.1: Querying the user with compatible input data

- Register: Users will enter information about the account they want to create, including their email address, their role (student and teacher), first and last name, and finally a password. The system will also force the user to confirm the password one more time, and the user's password is always encrypted. Besides, there will be a message sent to the email that the user registered, they must enter the correct OTP code to activate the new account, this helps improve the security of the connection between users and the system. Then, here is the way the system adds the users' data to the database.

```
result = QueryParamFunc('INSERT INTO "user" (id, email,
    password, firstname, lastname, role) VALUES (%s, %s, %s,
     %s, %s, %s)', (id, email, hashed_password, firstname,
    lastname, role))
```

Listing 3.2: Inserting the user with compatible input data

- Forgot password: Users will enter their information about their email address, and then, the system will check the existence of it. If the entered email is valid, a messenger will be sent to this email to confirm by OTP code. After that, users will be navigated to the reset password screen. Here is the code for updating the new password.

```
result = QueryParamFunc('UPDATE "user" SET password=%s WHERE
     email=%s',(hashed_password, email))
```

Listing 3.3: Updating the user with compatible input data

- CRUD actions:
  - Create a post: Users can create new posts on the forum. They enter the content, and title, and may add other related information.
  - Read all/a post(s) (including personal posts): Users can read their own and other people's posts. The interface displays a list of posts and content details when selected.
  - Update a post: Users can edit the content, title, or any information in their post. This function helps update information or correct errors if necessary.
  - Delete a post: Users can delete their post if they no longer want it displayed on the forum. This action requires confirmation to avoid accidentally deleting the post.
  - Create comments: Users can comment on their own or other people's posts, and the owner of the post will receive a notification when someone comments on their post.
  - Upvote a comment: If users find the comment useful, they can click on the heart icon to upvote this comment.

3. Database:

   PostgreSQL is a highly effective object-relational database management system. It is an open-source, cost-free application. Consequently, PostgreSQL may be utilized, modified, and distributed in any way. PostgreSQL's low maintenance requirements are a result of its exceptional stability. Therefore, the cost of ownership for PostgreSQL-based application development will be reduced in comparison to alternative data administration systems. PostgreSQL boasts a robust set of features

designed to cater to a wide range of users. Its flexibility lets developers make a wide range of apps, and its data integrity and fault tolerance features help managers keep things running smoothly. Whether data volumes are small or large, PostgreSQL provides a reliable and scalable platform. Additionally, its open-source feature and extensive extensibility further make its position as a powerful database solution. Among the basic data types, there exist integer data types such as integer and bigint, float data types such as real and double precision, and text data types such as character and varchar. In addition, PostgreSQL supports date and time data types such as date, time, and timestamp, which help manage time information flexibly. Boolean data type helps represent true/false values. This system also provides array data types such as integer[] and varchar[], which allow storing multiple values in a single column. JSON and JSONB data types json and jsonb support storing data as JSON. PostgreSQL also has a special data type like uuid to store UUIDs.

4. Libraries/Plugins:

   In addition to the technologies representing the three main components mentioned above, extensions and plugins play a very important role because they are effective support tools for specific purposes. Below, we will present a few important and most-used libraries in the system:

   (a) **Axios**:

   In the MVC model, the HTTP (Hypertext Transfer Protocol) protocol plays an important role in connecting and communicating between the main components of this architecture. Through these requests, the Controller can interact with the Model to perform CRUD operations, including creating, reading, updating, and deleting data. Specifically, for the calculation feature in the system, the user, standing on the Client side, will interact with the View to transfer data to the Server side via HTTP protocol for the Controller to process, after the calculation is completed, by Using the POST method, the Server side will respond with results and display them to the user. For the feature of displaying posts on the forum, the Controller is responsible for querying data from the database, then through the GET method, the Client side can receive the desired data.

   And to be able to do this, the system uses Axios, which is a JavaScript library used to make HTTP requests. Axios provides a convenient and easy-to-use way to interact with APIs and servers. This library provides simple methods such as GET, POST, PUT, and DELETE to make HTTP requests, as well as support for automatic data conversion between JSON and JavaScript Object formats. Axios also focuses on convenient error handling, providing easy control over request sending and response handling from the server. The example below is how Axios sends a GET request to a Flask endpoint to get data from the server:

```
1 axios.get('/api/data')
2   .then(response => {
3     // Handling the received data from server
4     console.log(response.data);
5   })
6   .catch(error => {
7     // Handling the error, if it exists
```

```
    console.error('Error:', error);
  });
```

Listing 3.4: Axios example

```
from flask import Flask, jsonify

app = Flask(__name__)

@app.route('/api/data', methods=['GET'])
def get_data():
    # T h c  h i n  logic v  t r  v  d  l i u
    data = {'message': 'Hello from Flask!'}
    return jsonify(data)

if __name__ == '__main__':
    app.run(debug=True)
```

Listing 3.5: Server handler example

In the above example, Axios sends a GET request to the Flask server's "/api/-data" endpoint, and the Flask server returns a JSON response.

(b) **WebView - MathQuill - MathView**:

Because this is a math app, fully displaying the math symbols is a must. The challenge here is that what has to be displayed on the user interface is the mathematical characters, while on the server side, to make the computing system understand the user's input, it is necessary for a different format, which is a latex string. Therefore the problem here is the need for a tool that acts as a converter from mathematical characters to latex strings and vice versa.

However, it doesn't stop there, converting latex strings into mathematical characters is quite an easy task, as there are quite a few support libraries, but there is a more complicated problem. Because this is a calculation application, users will enter their math problems to send to the system, so the character input tool can dynamically convert from latex to math symbols, as well as interact with the mathematical characters such as editing, and deleting while still maintaining the correctness of the latex string are much more challenging because for mobile platforms, up to now there are no libraries or plugins that can assist with this.

After our research, we decided to use MathQuill, which is a JavaScript library that allows users to enter mathematical formulas directly. MathQuill provides an interactive user interface for entering mathematical formulas conveniently and intuitively. However, MathQuill can only work on the Web platform, so WebView was used as a component that allows integrating a web browser into mobile applications. Through WebView, it is possible to display web content, interact with websites, or integrate existing websites into the application. And because WebView only allows programmers to use vanilla JavaScript for development, building a keyboard for user input will be more difficult.

```
const Fundamental = ({navigation}) => {
  const webviewRef = useRef();
  const handleSubmit = data => {
    const getData = async () => {
      try {
```

```
6         await axios
7           .post('http://localhost:8081/Calculus/fundamental', {
              data}) //here
8           .then(res => {
9             if (res.data.message) {
10              console.log(res.data.message);
11            } else {
12              navigation.navigate('Fundamental SOL', {
13                data: res.data.result,
14                equation: res.data.equation,
15                step: res.data.step,
16                img: res.data.img,
17              });
18            }
19          });
20      } catch (error) {
21        navigation.navigate('TabNavigator');
22      }
23    };
24    getData();
25  };
26
27  function onMessage(data) {
28    handleSubmit(data.nativeEvent.data);
29  }
30
31  return (
32    <SafeAreaView style={{flex: 1}}>
33      <Header nav={'CalculusList'} />
34      <WebView
35        ref={webviewRef}
36        style={{flex: 1, padding: 0, margin: 0}}
37        scalesPageToFit={false}
38        bounces={false}
39        scrollEnabled={false}
40        mixedContentMode="compatibility"
41        onMessage={onMessage}
42        onError={() => console.log('Something went wrong')}
43        source={{
44          html: html, //MathQuill HTML file
45        }}
46      />
47    </SafeAreaView>
48  );
49};
50
51export default Fundamental;
```

Listing 3.6: Interaction of the WebView to display HTML content

The above code can be understood as follows:

- Use useRef to create a reference to the WebView, making it possible to interact with the WebView from the React Native source code.
- The "handleSubmit" function makes an HTTP POST request to a specific address, sends data from the WebView to the server and processes the returned results.
- The "onMessage" function is triggered when there is an onMessage event from the WebView, this function processes the received data and calls the

handleSubmit function. WebView is used to display HTML content and interact with the WebView through events such as onMessage.

- The results from the server after the HTTP request are processed, and if valid, the program redirects to the 'Solutions' screen with the corresponding information.

After the system calculates the result, the returned data will be formatted as a latex string. Instead of displaying this string of characters to the user, which would be confusing, the system uses React Native's "react-native-math-view" library to convert the output latex string into mathematical characters. It is the entire process of transmitting, receiving, calculating, and displaying the final solutions to the user when they choose to enter manually from the system's keyboard.

(c) **Redux**:

Redux is a state management tool. Although it is used primarily with ReactJS, it can be used with any other JavaScript framework or library. It's very light at 2KB (including dependencies). The main reason lead us to apply this tool is that Redux helps manage the global state, keeping data accessible and updated without having to go through multiple levels. This increases performance and reduces complexity when working with the application state. Simultaneously, Redux provides an organized structure for the application, which will make the app expansion and maintenance easy. It is vital due to the various actions in community functionality.

Redux includes some main components, such as:

- Store: is where the application's state is stored. It maintains a state object and application components can access this state to read information or change the state by sending actions.
- Action: will describe the mutate in the app's state. It is necessary to declare the type of each action for the classification.
- Reducer: plays the responsibility to handle the action, and then returns the new value for the state.
- Middleware: is the layer between Reducers and Dispatch Actions. The place where Middleware operates is before Reducers receive Actions and after an Action is dispatched().
- Dispatch: is a store method, used to dispatch actions to reducers. When an action is passed, it triggers reducers to update the state.

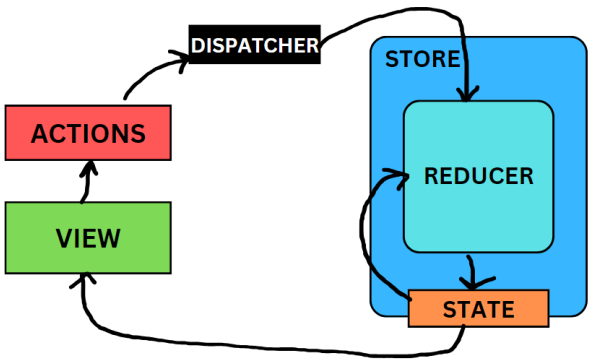


Figure 3.9: Redux workflow

(d) **latex2sympy2**:

This is a library that helps convert latex code to Sympy. This makes calculations easier and faster. The strength of the library is that besides converting latex string to sympy format, it can calculate directly based on the converted results and return relatively complete and accurate results. This saves many steps, making programming faster and the code file structure easier to maintain. The below block of code displays the way to utilize this library:

```
from latex2sympy2 import latex2sympy, latex2latex

tex = r"\frac{d}{dx}(x^{2}+x)"
latex2sympy(tex)
# => "Derivative(x**2 + x, x)"
latex2latex(tex)
# => "2 x + 1"
```

Listing 3.7: latex2sympy2 implementation code

(e) **Sympy & Numpy**:

Sympy is an open-source Python library designed to support symbolic computations. Sympy's goal is to provide a solution for mathematical and symbolic calculations, the same way we perform them on paper. Math operations can be performed, from basic operator calculations to more complex things like solving equations, algebra, integration, differentiation, and many other symbolic calculations. Because the goal of the project is not only to show the calculation results but also how to do it. Using Sympy, you can analyze equations by distributing, setting common factors, or implementing equality constants, this helps the method become more specific and easier to understand. For details:

```
from sympy import symbols, factor

# Define a symbolic variable
x = symbols('x')

# Define an expression
expression = x**2 - 4

# Factorize the expression
factored_expression = factor(expression)

# Print the original and factored expressions
print("Original expression:", expression)
print("Factored expression:", factored_expression)
```

Listing 3.8: Using Sympy for factoring the equation

The result should return:

```
Original expression: x**2 - 4
Factored expression: (x - 2)*(x + 2)
```

Listing 3.9: The factored equation result

In this instance, the phrase $x^2 - 4$ and the symbolic variable x are defined. The expression is then factorized using the factor function. The equations is factored into $(x-2)(x+2)$

For scientific computations in Python, NumPy is the foundational package. The Python library in question offers an array with multiple dimensions, as well as a variety of derived objects (including masked arrays and matrices) and an extensive collection of operations (mathematical, logical, shape manipulation, sorting, selecting, I/O, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation, and more) that can be executed rapidly on arrays. Numpy can be combined with Sympy to support algebraic equation analysis, however, it is also used in preprocessing input images for AI models, this will be clarified later.

(f) **OpenCV**:

OpenCV, which stands for Open Source Computer Vision Library. This library provides numerous powerful functions and tools to perform many tasks related to image processing and computer vision. There are some methods for preprocessing images:

- cv2.cvtColor(): Converts color space between different representations like RGB, BGR, HSV, Grayscale, etc.
- cv2.GaussianBlur(): Apply a Gaussian filter to blur the image.
- cv2.Canny(): Uses the Canny method to detect edges.
- cv2.HoughLines(): Detect straight lines using Hough transform.
- cv2.findContours(): Find contours in the image.

3.1.3 Testing and Deploy

For convenience, this application has been used in the Android Studio, which has a responsibility for providing an emulated mobile device platform. However, to reach the end user, the software must be allowed to run on a real device. Therefore, deploying the software on a mobile device to test the functionality of the system is paramount.
To display the entire software on a real device, connecting via wifi is an effective and simple way:

- Running "gradlew assemble" in "android" inside the React Native folder to export the APK file.
- Run this apk file on the mobile device.
- Go to Dev Settings → Debug server host port for the device to change the machine's IP, which can be checked by "ipconfig" in the command box.
- On the server side, replace the current URL with this IP with port 8081.
- Reload the app.

In addition, to reach end users, we have a test version deployed on the Android platform, which can be downloaded from Google Play. Google Play provides safe, reliable applications, users can trust that downloading applications from CH Play is safe and does not encounter malware or malware problems.

3.2 AI model implementation

3.2.1 Hybrid-ViT architecture

Hybrid Vision Transformer (Hybrid-ViT) combines convolutional neural network (CNN) and Vision Transformer (ViT) components for image processing (40). In the Hybrid-ViT model, the first CNN layers are used to extract features from images. These features are then fed into a ViT network, and the following process is similar to the above-mentioned. Hybrid-ViT aims to take advantage of CNN in local feature extraction. This improves the model's performance in many tasks, including optical character recognition. There are reasons why this model is more optimal for the OCR problem than the traditional ViT model, specifically:

- Although ViT is very effective in modeling complex and comprehensive relationships in image data, it is not always effective in extracting local features - an important factor in symbol recognition. CNN, on the other hand, is very powerful at extracting local features such as edges and corners that are necessary for character recognition. This combination in Hybrid-ViT ensures that the model can understand both the general and local details. Hence, the accuracy of the model will achieve a significant improvement.

- OCR often has to deal with noisy or distorted images. CNN in Hybrid-ViT helps the model resist noise and recognize characters in image variation conditions, while ViT contributes to analyzing the relationship between characters and the context around them. Therefore, in this section, each part of the model will be analyzed.

1. **Backbone**

   Pix2tex uses ResNetV2 as the backbone layer for the Hybrid-ViT model. This model is a refined model based on the RestNet model, so they have many similarities with each other. The ResNetV2 model is divided into stem and stages. The stem is the first part of the network, right after the input layer. Stem is responsible for processing the initial input data and preparing it for the next parts of the network. Stem included:

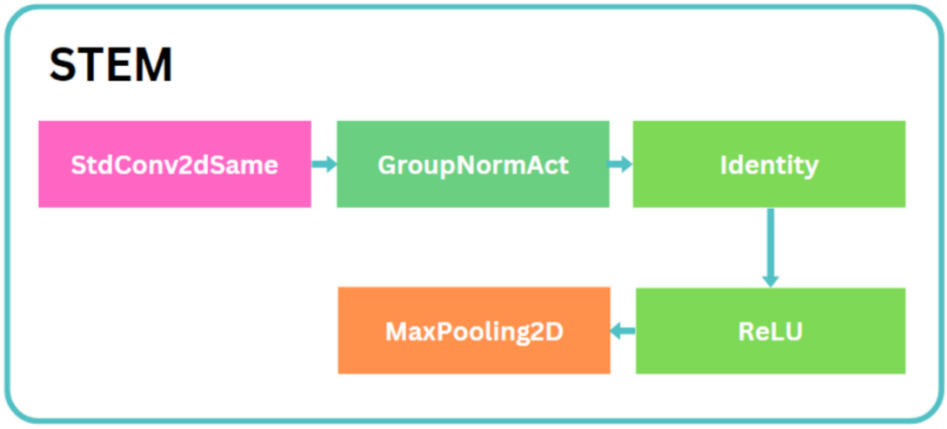


Figure 3.10: ResNetV2 architecture

- **StdConv2dSame**: This is a two-dimensional convolution layer. It uses a kernel to extract features from the input data. This entails computing the

weighted sum of pixels in a neighborhood defined by the filter's size [50]. Padding is also used to ensure that the output's spatial dimensions (width and height) remain consistent with the input data. The image below shows the way it works:

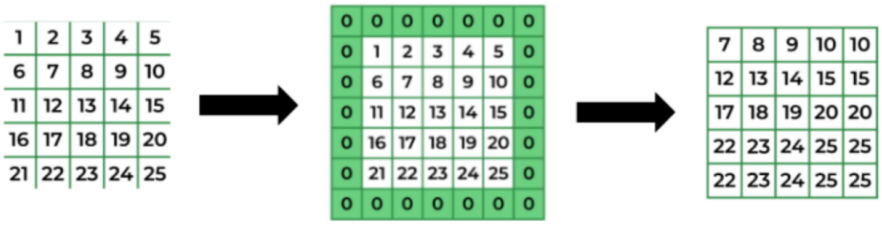


Figure 3.11: Kernelizing method of StdConv2dSame layer

- **GroupNormAct**: GroupNormAct is a combination of Group Normalization (GroupNorm) and activation function (Act). Group Normalization is a normalization method used in convolutional neural networks. Unlike Batch Normalization, Group Normalization does not depend on the size of the batch. Instead, it divides the output channels of a convolutional layer into groups and normalizes by calculating the mean and variance. The formula here is:

$$u_i = \frac{1}{m}\sum_{k\epsilon S_i} x_k, \sigma_i = \sqrt{\frac{1}{m}\sum_{k\epsilon S_i}(x_k - \mu_i)^2 + \varepsilon}$$

- **Activation function**: The activation function is an important component in a neural network, helping the model learn non-linear relationships. The activation function in this model includes ReLU and Identity. ReLU is a popular activation function in neural networks. Its formula is $f(x) = max(0, x)$. It can be understood that if the input x is positive, it will return that value, and if x is negative, it will return 0. While the Identity Function is a simple function, where the output is exactly equal to the input, $f(x) = x$.
- **Maxpooling2d**: it is used to reduce the size of feature maps after going through convolution layers.

This architecture also has a concept of "ResNet Stage", which describes specific parts or layers of the model. Each stage consists of a series of convolutional and normalization layers stacked on top of each other and often includes layers of increasing complexity. In the ResNet Stage, this Bottleneck is an important network component, applied to reduce computational complexity and increase model performance. The number of bottleneck layers is based on the number of backbone layers, it includes many main classes, which are displayed below:

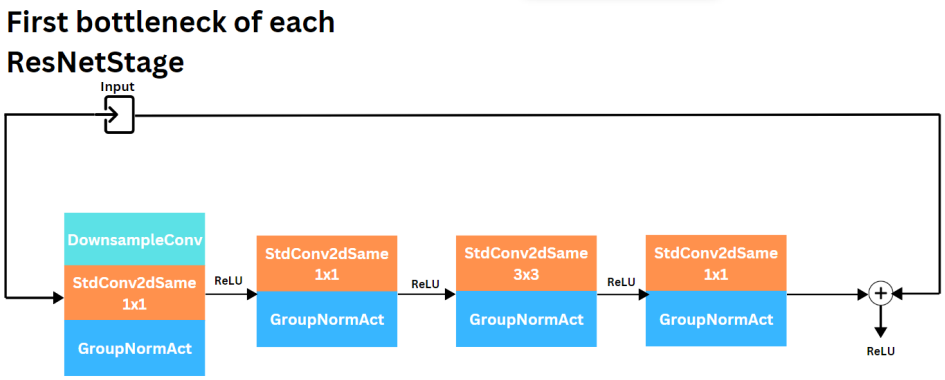


Figure 3.12: First bottleneck of each ResNetStage

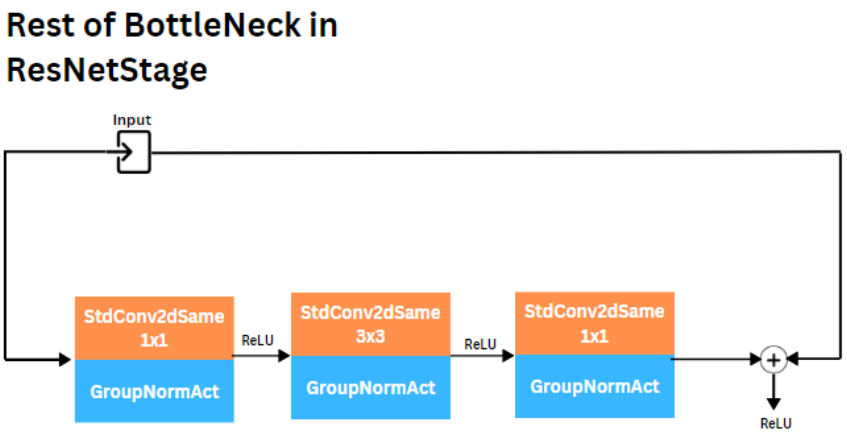


Figure 3.13: The rest of bottleneck of each ResNetStage

- The DownsampleConv class in the bottleneck of the ResNet model has the goal of reducing the size of the feature map, while the other convolution layers in the bottleneck reduce and increase the depth of the feature map without changing the size. When data passes through a Bottleneck block, its spatial dimensions (width and height) may need to be reduced to enhance higher-level feature extraction and reduce the amount of computation.
- First 1x1 Convolution layer (diminishing): This layer reduces the number of input channels, reducing the depth of data. This reduction in depth reduces computational costs.
- Central 3x3 Convolution layer (bottleneck): This layer processes the input after it has been reduced in depth by the 1x1 convolution layer. 3x3 convolution layers help learn complex features from data.
- Final 1x1 Convolution layer (expanding): This layer increases the depth of the data before adding it to the original input. It ensures information reduced in depth in the first step can be effectively added back into the data.

According to the project configuration, $LayerBackbone = [2, 3, 7]$, the model will have 3 stages, and each stage will correspond to the value in the array. The first stage will have a total of 2 bottleneck layers, so there will be a total of 6 convolutional layers and 6 norm layers. The second stage has 3 bottleneck layers, with 9

convolutional layers and 9 norm layers. The final stage has 7 convolution layers, so there will be 21 convolution layers and 21 norm layers. It can be seen that a larger number of bottleneck layers can help the model learn more complex features from the data.

2. **ViT model**

   Due to the characteristics of this Hybrid-ViT model, it is essential to incorporate the vision transformer. The previous section states that the vision transformer model will be implemented sequentially. However, due to the purpose of this model being used for optical character recognition, variations will be different from a typical classification problem. Specifically, with image classification models, the ViT model will have a classification layer at the end to determine the type of image, such as garbage classification or animal classification. However, with the OCR problem, classification is not necessary, because the output will usually be a string corresponding to the image content, so in this case, using a Transformer Decoder will be much more appropriate. Therefore, the complete process of the Hybrid-ViT model for the problem of recognizing characters, equations, and mathematical formulas will be presented below.

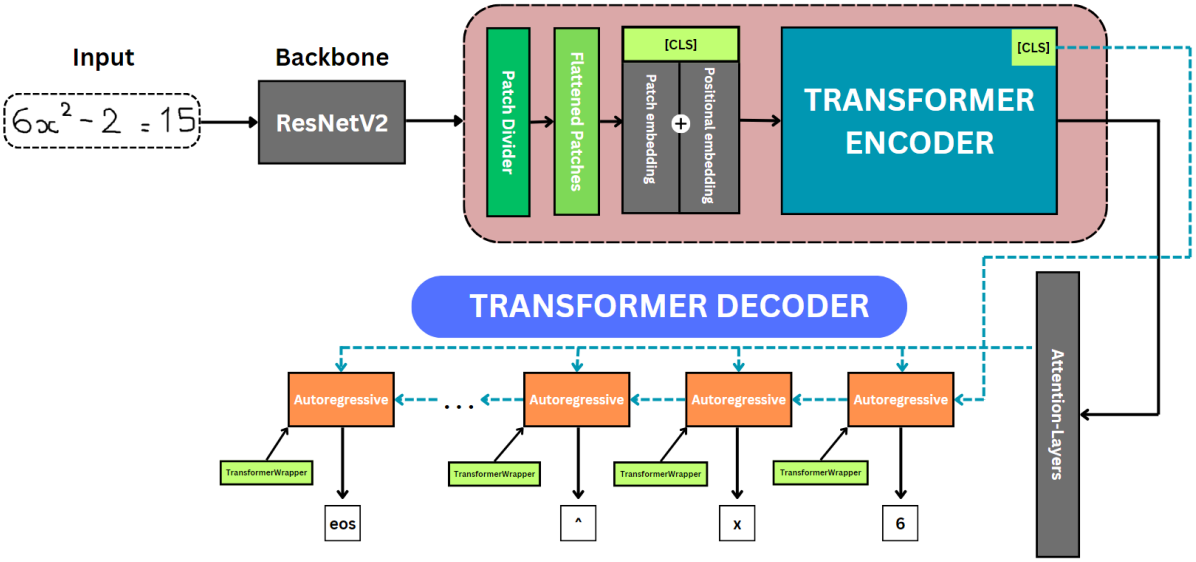


Figure 3.14: Full Hybrid-ViT architecture

- The process shown above consists of three main parts, first is the backbone, second is the transformer encoder, and finally the transformer decoder. The transformer decoder plays an important role in predicting the output sequence of the problem by learning the relationships between characters. About decoder, is an important component responsible for converting the encoder's output into the final output data that the model needs to predict or generate. The decoder in Transformer is widely used in applications such as automatic machine translation, text generation, and many other natural language processing (NLP) tasks. While it is frequently employed in NLP models, OCR problems, particularly those involving character recognition and mathematical equations, necessitate its inclusion. The desired outcome of the model is a latex string that accurately represents the user's concern. Each mathematical sign depicted in the image is associated with a certain syntax in LaTeX. Therefore,

the model must possess the ability to forecast the syntax based on the acquired data. Hence, opting for a transformer decoder is a logical decision.

- Decoders, also known as autoregressive transformer models, are trained on the traditional language modeling issue of predicting the next token after reading the ones that came before. They correspond to the decoder in the original transformer model, and a mask is placed on the entire phrase so that the attention heads may only sense what has come before in the text, not what has come after. Although these models can be fine-tuned to generate good results for a wide range of jobs, text generation is the most natural use.
- ViT introduces an extra token called x-class, which serves as a learnable embedding. This token, inspired by the concept used in BERT, is referred to as the [CLS] token. In the training process, the [CLS] token is combined with other spatial tokens in the current patch embeddings. Through a self-attention mechanism, information is compelled to flow from all other tokens to the [CLS] token [48]. Consequently, the embedding of the [CLS] token can be regarded as a comprehensive representation of image features. It can be utilized as an initial hidden state for the model's decoder, instead of using the entire encoder's output feature maps. So it can be understood that this token represents the first character in the string that will be predicted
- The feature vectors from ViT will serve as input to the Transformer decoder. These features contain information about the image that the decoder will use to generate the character string.
- In the autoregressive process, the decoder generates tokens one by one, each time based on the sequence of previously generated tokens and feature vectors from ViT. The attention mechanism is used to determine the relationship between the generated tokens and the feature vectors [51]. For details.
  - Self-Attention Layer: Allows the model to consider every position in the input sequence to produce an output. This mechanism is often implemented with multiple attention "heads" running in parallel, allowing the model to focus on many different relationships at the same time, called Multi-head Attention. It creates a context for each token generation step, helping the model understand the relationships between tokens and generate more accurate predictions.
  - Feed-forward Neural Network: Linear neural network applied after self-attention, usually consisting of two layers with a non-linear activation function in between. The feed-forward layer plays an important role in processing and refining information after it has passed through the attention mechanism. Each unit in the feed-forward layer processes information from a specific position in the input chain independently of other positions. That is, this network does not share information between different locations in the chain. Additionally, it helps the model learn more complex representations, based not only on relationships between tokens but also on the unique characteristics of each token.
- After that, a linear layer converts the output vector from the last transformer block into a vector of size equal to the vocabulary size [51]. Then, the softmax layer converts this vector into a probability distribution over all possible tokens in the vocabulary. Specially, the model will choose the next token with the

highest probability. Here, the model uses top-k, with k = 5, which will help retrieve 5 values with the highest probability distribution, to make the final prediction for the character [52].

- One issue to note is that when training a Transformer model to generate sequences, we often feed the entire desired output sequence to the model at once to optimize computational efficiency. However, we don't want the model to "cheat" by using information from later tokens in the chain to predict the current token. Therefore, the model applies an additional attention mechanism, called Masked attention, to prevent the model from using the data of the following characters to guess the current character. Thus, the model will accurately reflect its performance. It can be seen that the last character of the string is always [eos]. When the model predicts that the EOS token has the highest probability, it will generate this token and end the chain generation process. The eos token plays an important role in specifying the end of a chain, preventing the problem of infinite word generation, in addition to helping the model not have to process unnecessary additional information after the chain has been completed.

3.2.2 Dataset

As mentioned, the AI model is built based on the Hybrid-ViT architecture, so using a rich and diverse data set for training is very important because a quality dataset will ensure high performance for OCR modeling. Using text images from many different writing styles, formats, and strokes helps the learning model to be diverse and flexible.
Initially, the pre-trained model pix2tex uses the im2latex-100k dataset, an already-built dataset for the image-2-latex system employment on OpenAI, over 100,000 formulas and pictures are divided into test, validation, and train sets [46]. This data set is considered large enough for a mathematical character recognition model because it includes many images from basic equations and formulas to more advanced and complex things. This ensures the model is capable of learning images of uncommon mathematical symbols. However, the weakness of this dataset is that it only focuses on printed characters without images of handwritten equations and formulas. This causes the pretrained pix2tex model to only be able to recognize equations imported from a computer. Even with a monotonous format, Lukas Bletcher's pix2tex model has difficulty recognizing distorted images and not in line. To be able to deploy this model to serve a mathematical application, where the input is diverse images of the user.

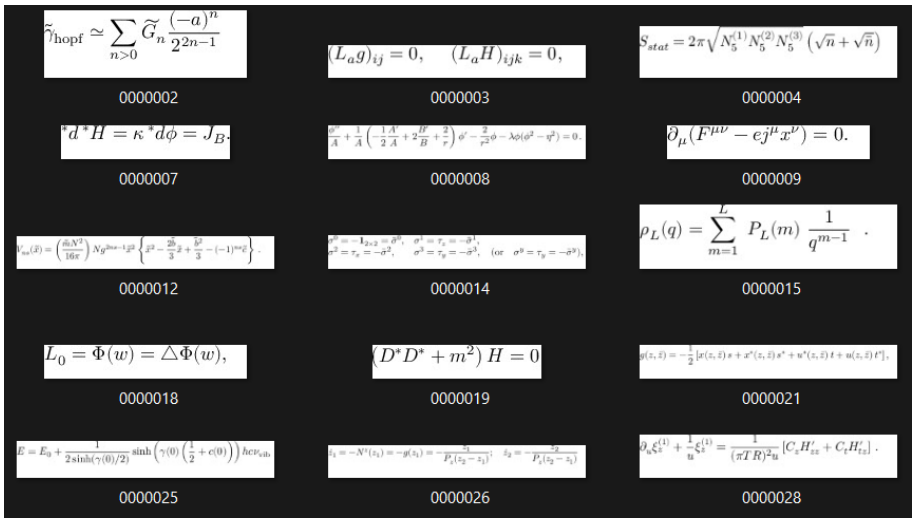


Figure 3.15: Images about the im2latex-100k dataset

Besides, it focuses on identifying common math problems to support students. So we combined the original pix2tex dataset with other datasets, including:

- Linear Equation Images dataset of Archan Ghosh on Kaggle [53]: The dataset provides a folder containing 24,000 images of randomly generated linear equations. In addition, all labels are stored in a CSV file. So to use the data, we need to use Python's data processing and analysis libraries, pandas, to map images and labels together.

- Images of Randomly Generated Quadratic Equations dataset of Archan Ghosh on Kaggle [54]: The dataset consists of 12,000 randomly generated images of quadratic equations. And like the dataset mentioned above, this dataset is also processed using the pandas library to map images and labels.

- CROHME (Competition on Recognition of Online Handwritten Mathematical formulas) dataset, which includes information on online handwriting of mathematical formulas [55]. By doing this, it is ensured that the dataset contains a variety of mathematical expressions that have practical applications. There are letters, numbers, and other mathematical symbols. Specifically, we use the dataset extracted from the directory of 2014, 2016 and a pre-labeled dataset with an unknown year, all consisting of 20,000 images. Because these datasets were not fully processed, we hand-processed several hundred images to ensure the data mapped together properly.

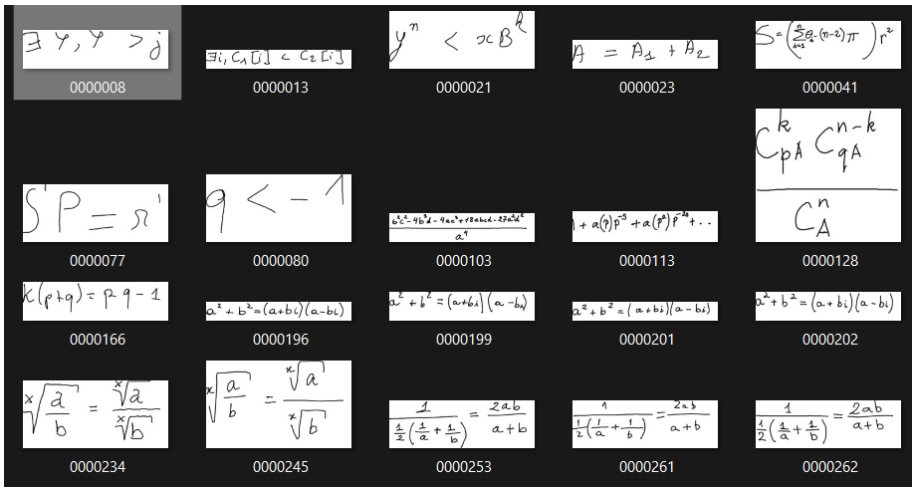


Figure 3.16: Images about the CROHME dataset

Input image from user about equations and formulas into a cloud-based platform. We use Cloudinary, which is a cloud-based service, that provides an image management solution including upload, storage, manipulation, optimization, and delivery. To collect more data, these images will be selected and aggregated to make images for the model to continue learning. And of course, this requires permission from the user.

3.2.3 Pre-processing image

To improve the accuracy of recognition results, pre-processing input images from the user is extremely necessary. To improve the accuracy of recognition results, pre-processing input images from the user is extremely necessary. Specifically, the pre-processing techniques applied to this model are done to reduce the complexity of the input image, making the model easier to recognize. In addition, there will be some cases where the model incorrectly recognizes an image when it has a few wrong characters in the string, so processing the output is also something that needs to be done, this will be presented more clearly later.

- To simplify the background of an image, we used the technique of converting image pixels into binary, in other words converting a color image into a black-and-white image, so the model will be able to recognize the image. It's easier when the data set to be trained is mainly black and white images.

```
1 _, thresholded_image = cv2.threshold(image, 128, 255, cv2.
    THRESH_BINARY)
2 _, img_encoded = cv2.imencode('.png', thresholded_image)
```

Listing 3.10: Using OpenCV for converting RBG image the black-and-white

The result should be.

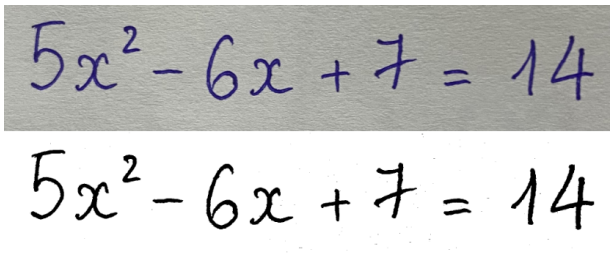


Figure 3.17: Binary converter result

However, this method also has some limitations, such as for images that already have a simple background, using binary conversion can cause the image to lose some important information, here is an example.

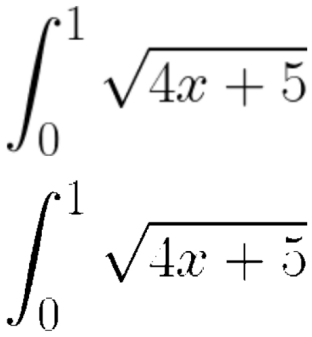


Figure 3.18: Binary converter can make noise to the image

- To overcome this drawback, we have designed a small algorithm to distinguish between black and white images and color images, with the purpose of only processing color images and noisy images to preserve the input images. Here's how to implement this.

```
1 white_pixels = np.sum(image == 255)
2 black_pixels = np.sum(image == 0)
3
4 white_black_ratio = white_pixels / (black_pixels + 1e-5)  # Tr nh
    chia cho 0
5 if white_black_ratio > 15:
6     print("Only-black-and-white image detection.")
7 else:
8     print("RBG image or noises detection")
```

Listing 3.11: Determining RBG and only-black-and-white image

  - The first is to use NumPy to count the number of pixels with values 255 (white) and 0 (black) in the image.
  - Then calculate the White/Black ratio and print it to the screen. Add 1e-5 to the pattern to avoid dividing by zero.
  - Finally, based on the value of the White/Black ratio, determine whether the image has a white background with black text or not.

- Ouput-processing: It is a certain fact that the model cannot 100% accurately recognize every input image for many reasons. The mistake here is incorrect recognition

of single characters inside equations and formulas because, in latex, there are many ways to display a character, such as with the character "x", which can be written into(`\textbf{x}`) or (`\mathbf{x}`), or depending on the user's handwriting, each person has a different way of writing, which the system can understand as the character χ (`\chi`). The technique here that can be applied is to use regex to reformat the latex string correctly. Below is an example of converting (`\textbf{x}`) to x.

```
def remove_latex_formatting(input_str):
    result_str = re.sub(r'\\textbf{([^}]*)}', r'\1', input_str)
    return result_str
```

Listing 3.12: Using regex to format the latex string

With the scope of student support, these complex characters will not be necessary. Therefore, to optimize the user experience, the application not only provides a single result for the user but also provides more options if the user presses the "Re-generate" button. Thus, users will have more options for their problems.

The entire pre-processing process is shown in the image below.

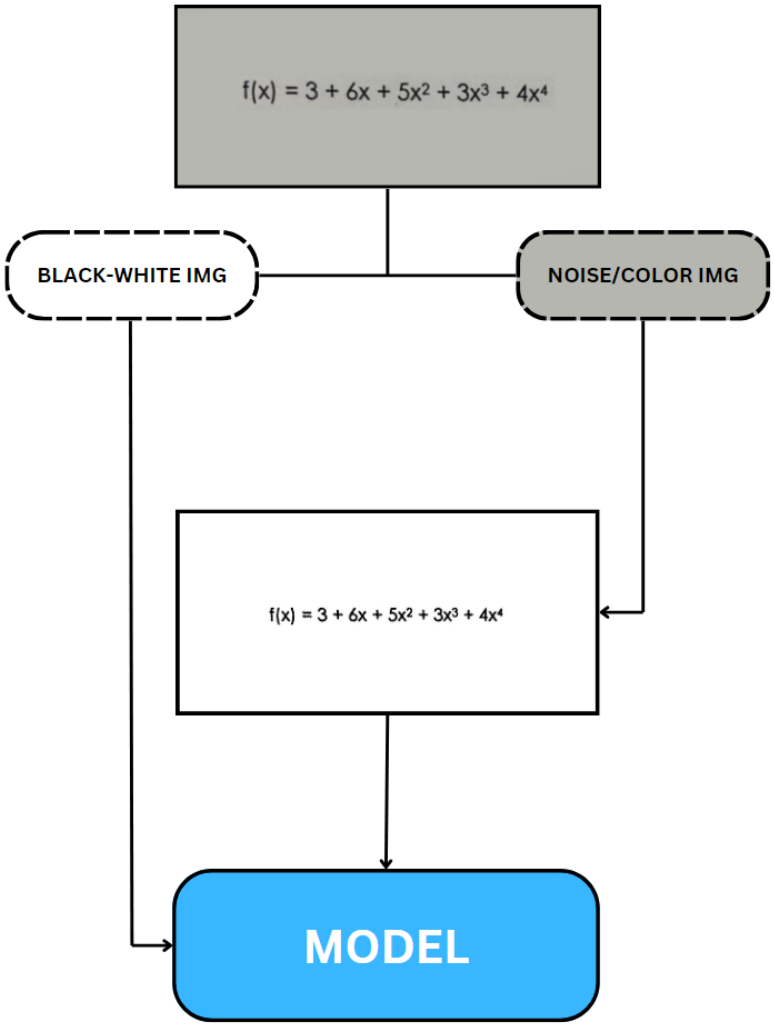


Figure 3.19: Pre-processing process diagram

Chapter 4

RESULTS

4.1 Mobile Application

4.1.1 Ability to solve problems

After part of it was completed and it would be possible to solve the math-related problems using basic mathematics operators, calculus, numerical methods, etc., AiMA may meet certain key values.

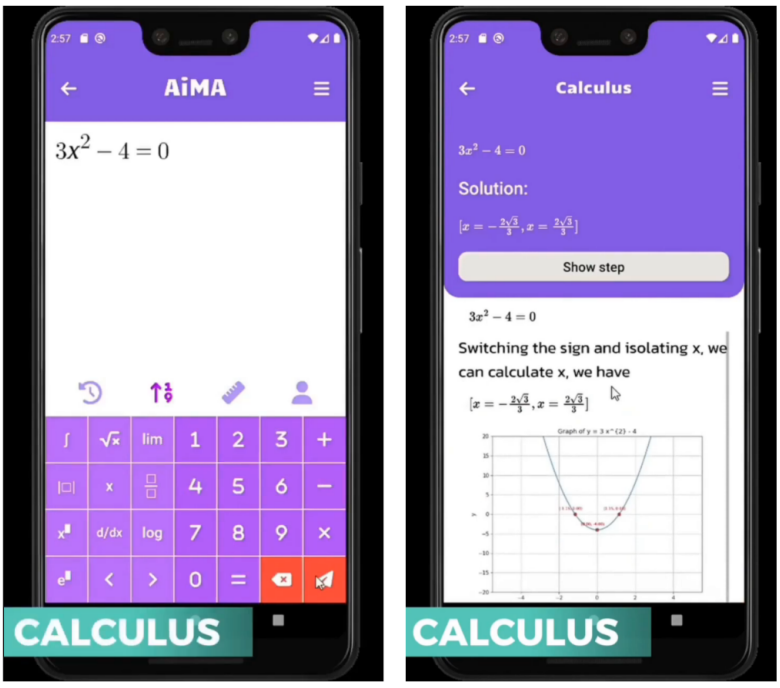


Figure 4.1: Algebra/Calculus interactive views

The image on the left is of the system's keyboard interface, users will directly enter data here to send their math problem to the system for processing, the image on the right is the result returned by the system, Includes answers and solutions along with function graphs.

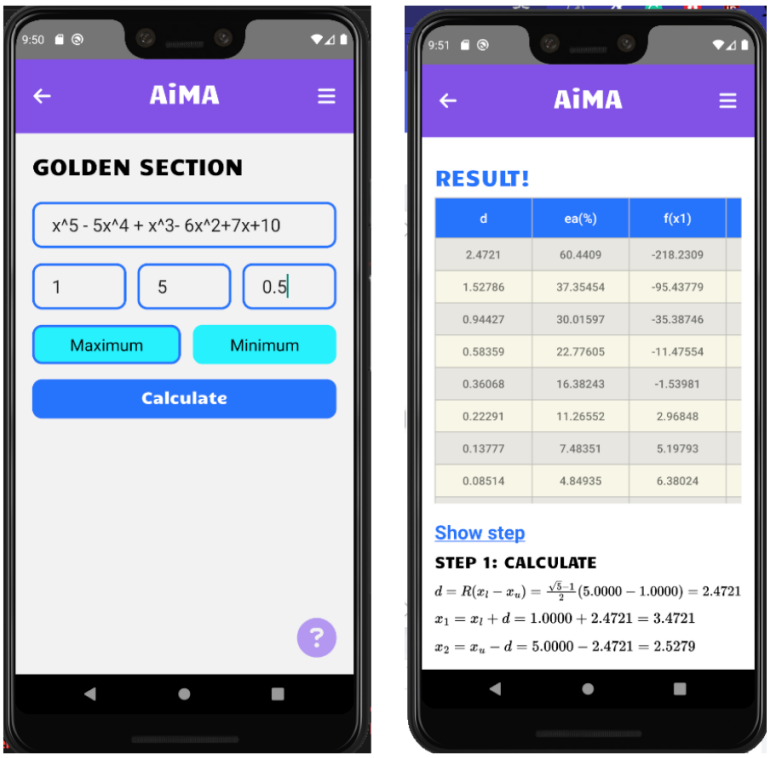


Figure 4.2: Numerical method (Golden Section Search) interactive views

Figure 4.2 shows how AiMA is shown when solving the golden section search optimization technique issue. This search application uses an iterative approach to identify the extreme point of the function rather than conventional mathematical methods like multi-level derivatives. The necessary data must be provided by the user. In this case study, for example, the function, the target error, the start and finish values of the given interval, and the condition about the minimum or maximum requirements are all strings. The request from the client is subsequently received by the server. Since this question will need iteration to be answered, choosing between a while-loop and a for-loop may be taken into consideration. Because a condition check is necessary after each iteration and the number of iterations to be performed is unknown in advance, the while loop is acceptable. Moreover, a while loop computation completes in only two hundred milliseconds as opposed to a for-loop computation's four hundred milliseconds. Because of this, it makes sense to compute the issues associated with the iterative technique using a while-loop. The user will, however, have to wait longer for a response from the server. Including the previously mentioned execution time, the server's response time to the client in this case study was two seconds. In this instance, the research was tested locally due to the server. When deploying the server in a production setting, there will be an increase in the measurement time due to network variables. Thanks to advances in client-side rendering performance and execution time optimization, this is fully manageable. Furthermore, after using AiMA to calculate the issue in this case study, the accuracy is comparatively good. When compared to confirmed manual calculations, AiMA yields results with a nearly 100% accuracy rate, and when compared to reliable computing applications, it is almost perfect. The returned numbers are rounded rather than shown in full to provide the user with the most logical view, thus even with faulty precision, it is acceptable. Currently, this mobile application for mathematics may be used on an Android smartphone. Specialized algorithms including Newton's method, Heun's technique, Euler's method, Simpson's rules, and the golden section search method are also updated regularly. They not only provide precise

results but the step-by-step box is also consistently presented in a clear, detailed manner. Performance may be significantly improved by using react hooks, such as useMemo and useCallback, which reduce the number of times the interface is displayed again.

4.1.2 Forum-related functionalities

Authentication

After implementation, the login and logout function has worked according to the set goals, as it requires the user many authentication steps, such as entering the OTP code, this ensures the security of the system.

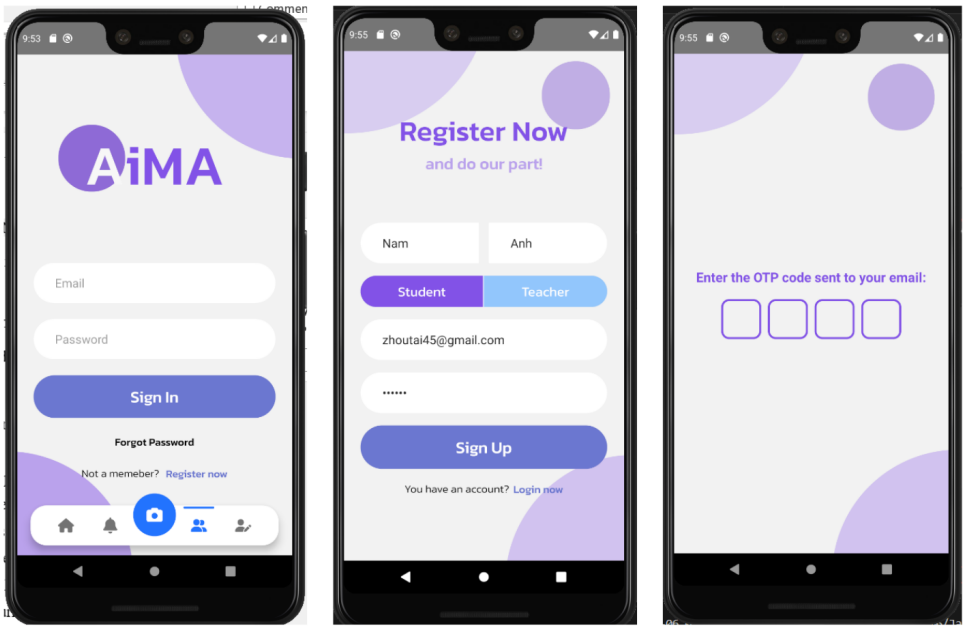


Figure 4.3: Login/Register/OTP view

The password recovery feature for users has also been fully implemented. By entering an email and authentication OTP code, users can reset the password for their account.

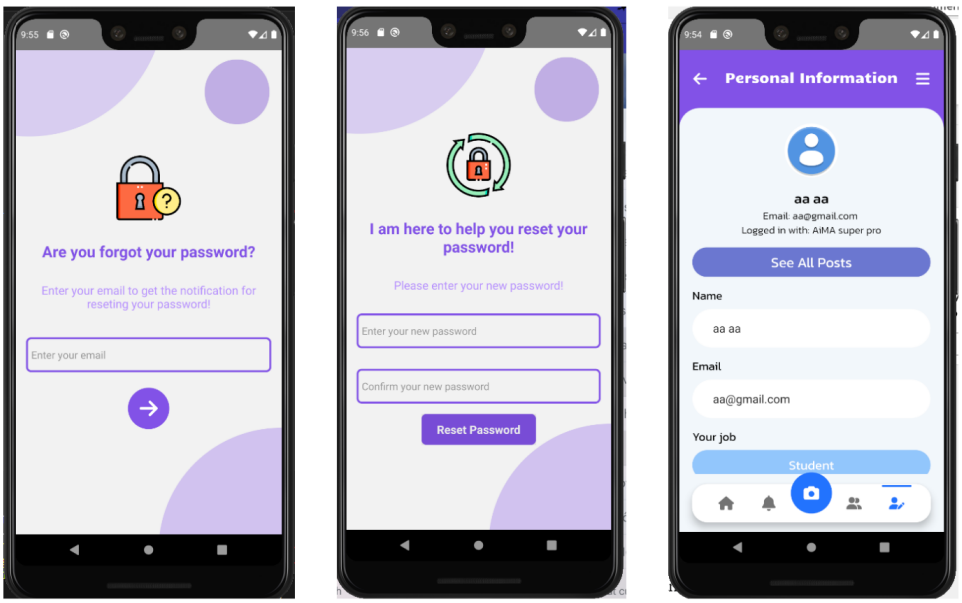


Figure 4.4: Confirm-email/Reset-password/OTP view

CRUD features have been completed, and users can easily interact with the system through interfaces as shown below.

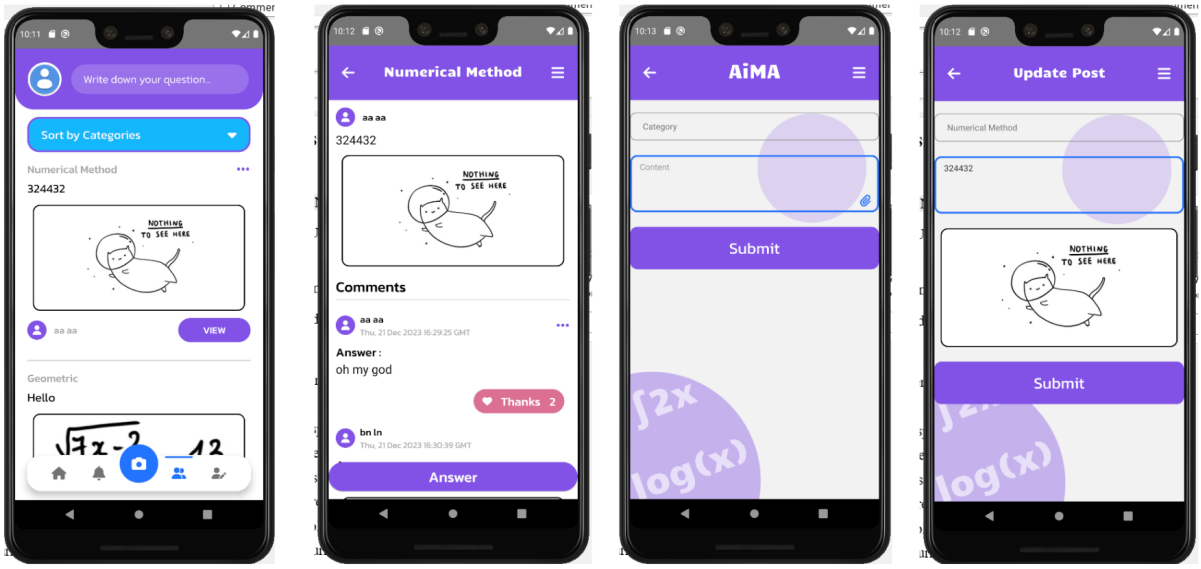


Figure 4.5: View-all-posts/View-a-post/Create-post/Update-post view

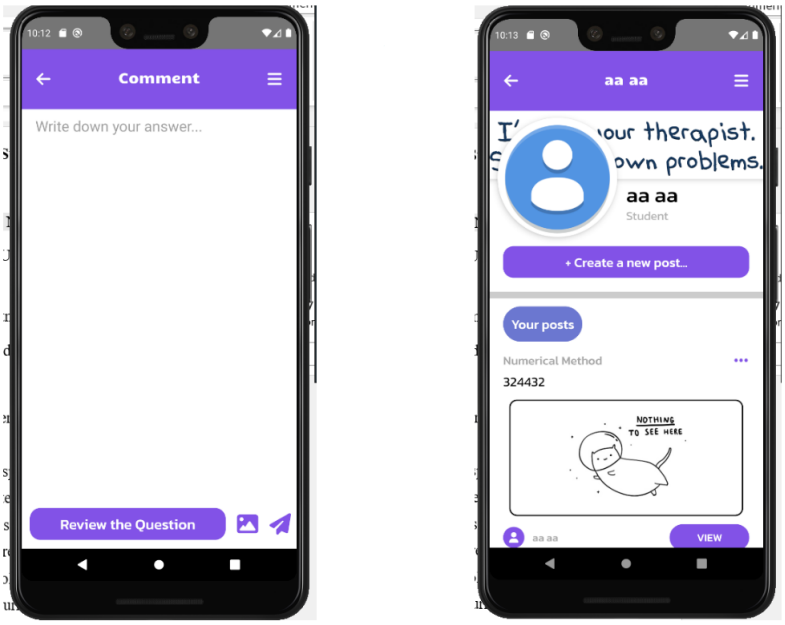


Figure 4.6: Comment-post/Personal view

The recently completed forum feature brings an important step forward for interaction and communication within the application. With the ability to discuss, share ideas, and solve problems, forums not only create a vibrant community but also enhance the user experience.

AI model

After deploying and fine-tuning the pix2tex model - a pre-trained mathematical model, we have achieved many positive results. The initial limitations of this pre-trained model have been significantly overcome. First, it is the ability to read users' handwritten images. As mentioned, this probability for the pix2tex model was virtually zero, and after fine-tuning the model was able to respond correctly. However, with handwriting, the current model will work best with problems and equations from basic to intermediate levels such as equations with steps, containing roots, fractions, etc. For more complex types of problems, with mathematical characters such as primitives, the model encounters a bit of difficulty as it cannot recognize 100% accurately.
In addition, the ability to identify images that are not aligned or have noisy backgrounds has also been improved. Previously, the model could only recognize images with white backgrounds and black letters, but now it can recognize more diverse backgrounds. However, if the background is too complex, such as a math problem written on a gridded page, the model may misidentify it.
Besides, the model can read more complex problems, such as limits, for which previous models had difficulty giving accurate results.
To be able to evaluate the model, we used wandb, a platform that provides services and tools to track, manage, and understand machine learning models and artificial intelligence projects. We will evaluate based on 2 criteria, token accuracy and BLEU score. Moreover, to start the training process, we use the GPU of a personal computer to train the model. The hardware index for this process is shown in the table below.

Table 4.1: Hardware and requirement details.

Windows/Ubuntu version	Window 11
CPU	12^{th} Gen Intel(R) Core(TM) i7-12700H (20 CPUs), 2.3GHz
RAM	16384 MB;
GPU	NVIDIA GeForce RTX 4060 Laptop GPU
CUDA version	12.2
Programming language	Python 3.7.16

And we get the following results after training.

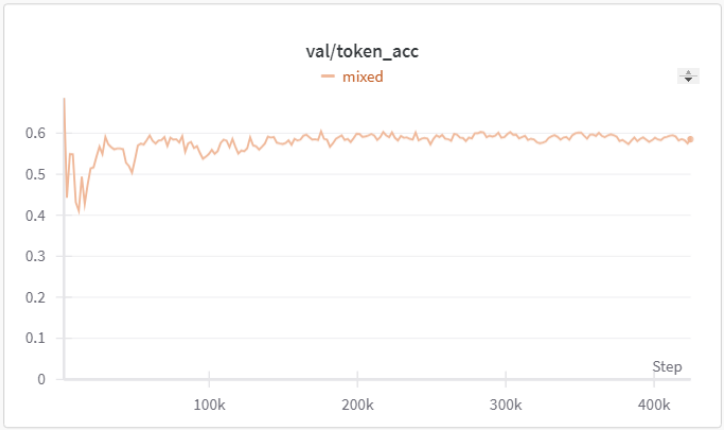


Figure 4.7: Token accuracy graph

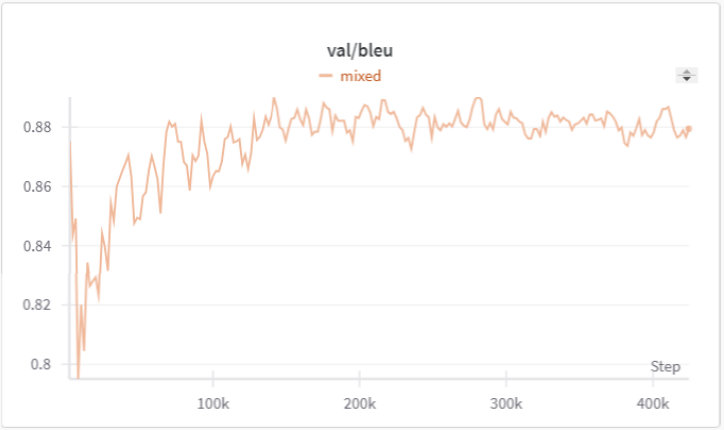


Figure 4.8: BLEU score graph

It can be seen that the model's accuracy is approximately 60%, while the BLEU score achieves nearly 90%. With this accuracy, it is relatively not high. This can be explained, because, for a mathematical equation represented by a latex string, there are many ways to write it, for example, with the character x, the string "x", `\mathbf{x}`, or `\textbf{x}`, or for trigonometric characters, "sin" and "cos", which can also be written as "`\operatorname{sin}`", "`\operatorname{cos}`", instead of "`\sin`", "`\cos`", respectively. In terms of display and meaning, they are the same, however, when compared

with sample labels, it can affect the evaluation process. Therefore, we additionally use the BLEU score to evaluate, with 90% accuracy, that it reflects the similarity between two character strings.
The above explanation is completely well-founded, when another model, using the same architecture, is Anh Duy Le's model, with an accuracy of 48.39% while BLEU score they achieved with 89.94% (ref), has shown a very positive result when being able to identify mathematical equations within their dataset. The hardware index for this process is shown in the table below.

4.1.3 Evaluation

Finally, after the implementation process, we were able to evaluate our application against related studies. To turn AiMA into a complete, commercially viable application, we will create a table to compare with famous applications currently on the market, to get an overview of continued improvement. The comparisons among AiMA and related mobile systems are shown in Table 1.

Function in general	Function in detail	AiMA	Photomath	Symbolab	Mathway
UI/UX	Easy to use	x	x	x	x
	Simple UI	x	x	x	x
	Smooth interaction	x	x	x	x
Response Time	Fast response	x	x	x	x
	Medium response				
	Low response				
Publication	Open-source	x			
Specialized Algorithms	Variety	x			x
Basic Algorithms	Complex equation	x		x	x
	Simple equation	x	x	x	x
AI Scanner	Gallery picker	x	x		x
	Camera scaning	x	x	x	x
	Response in time	x	x	x	x
	Handwriting recognition	x	x	x	x
Result/Steps showing	Correctness	x	x	x	x
	Completeness	x	x	x	x

Table 4.2: Comparisons among AiMA and related mobile system

It can be seen that the proposed application has stronger points than other applications, such as the ability to solve more types of math aspects which is popular in the community of Vietnamese students, or an easier-to-use and friendlier interface. However, there are still some limitations, which will be mentioned more clearly in the next section.
Besides, because AI is the core technology of the system, it is useful to compare the proposed model with other models, including the model on which fine-tuning is based, to have an overview.

Criteria	Our model	pix2tex	CoMER	Khiem's model	Anh Duy Le's model
Handwritten recognition	x		x	x	
Color images recognition	x			x	
Various symbols recognition	x	x	x		x
Integrated products (math app)	x			x	
Ability to solve equations	x				x
Open-source	x	x	x	x	

Table 4.3: Comparison of Different Models

Chapter 5

Discussion, Conclusion, and Future Works

5.1 Discussion

The value and influence of mathematics in people's lives are increasing as technology develops. Particularly in education, the rise of mobile technologies has brought about several advantages, including portability, ease, and accessibility. Creating a mobile system that can serve college and university students and high school kids is essential, providing answers to simple math problems and more advanced algorithms. The mathematics mobile applications published on the market generally provided various features with a high accuracy rate calculator system. Nevertheless, these software products are only designed to support mathematics topics for 12th-grade students and below, which means that the shortage of advanced fields can limit their reachability to another group of customers, who are university students. In addition, despite the understanding that money is used for maintaining the application, from the perspective of someone who has experienced many mathematical applications, it is possible to realize the limitations of the free package of some applications, including Mathway. So to be able to create the best application and be able to reach the end user, the consideration of creating a product that is priced, even for free, is worth considering.

5.2 Conclusion

To meet this requirement, this study introduces AiMA, a mobile AI-based system that helps college students with math-related problems. The MVC model is used to host, deliver, and manage the resources and services that the client requests. When aiming to build a software product, it is essential to apply the development model—in this case, the waterfall methodology. This model is easy to manage due to its clean structure. Besides, the waterfall is the ideal software development model for small-scale software projects with specific requirements. The research mainly concentrates on specialized fields in mathematics, such as numerical methods in theoretical models in computing subjects, and probability & statistics problems. AiMA offers substantial advantages over competing alternatives by enabling users to solve challenging calculus problems and get customized answers for specific mathematics topics, according to an evaluation of the system's capabilities. AiMA is an invaluable mobile system for college and university students who need specialized assistance with their mathematics studies. Moreover, to

make the user experience better, AI models are used as the most optimal measure. With the available pre-trained model, we apply pre-processing methods, model refinement, as well as data diversification to improve performance, making this model more suitable for a specific application. use mathematics. And certainly, model development efforts do not stop here, because we recognize the remaining limitations. This will be made clearer in the next section of this chapter.

5.3 Future works

5.3.1 AI improvements

The goals set out in improving the model have been completed by about 90% because there are still limitations that are in the process of being overcome. For example, handwriting can be recognized, but with complex equations, the system will have difficulty. In addition, with images containing too complex backgrounds, it will also be difficult to identify. In the future, we will work to train the model continuously so that it can read and respond to a wider variety of images. For details, there are a few ways that can be used to improve the AI model, the first is to be able to collect more data, because the more data, the more accurate and efficient the AI model will be. Besides, enhancing the ability to clean and pre-process the data to reduce noise, outliers, and inconsistencies is crucial before training the model. Techniques like dimensionality reduction, feature scaling, and data standardization may be used. In addition, programming more specialized fields related to mathematics for the application: recognizing the complexity and variety of problems, and diversifying solutions is a necessity.

5.3.2 User interface improvements

Although the application's UI up to now has been suitable for most users. However, designing an interface that is not only user-friendly but also attractive is something that takes effort. Therefore, to commercialize this application, editing the interface as well as further optimizing the user experience is essential. This requires the research team to create an overall design for the software interface, the proposed application is Figma.

5.3.3 Instructions improvement

Regarding the field of algebra and calculus, in the future, we will continue to research to be able to provide many ways to solve many different types of problems, because currently, the system is only providing solutions for popular and fundamental math problems. Types of problems such as solving equations from 1st to 3rd degree, anti-derivative, and integral have been solved, but derivative, logarithm, and limit problems have not yet been completed. Therefore, the future goal is to be able to design specific steps for these types of math.

5.3.4 Chatbox development

Realizing the popularity of chatbox technology today, we plan to develop a chatbox capable of answering users' questions, not only about calculations but also related to mathematical theory. Thus, users can receive more comprehensive support. We recommend

reusing open-source models, such as llama2 for continued development, thus saving development time and costs.

Bibliography

[1] Rashedul Islam, Rofiqul Islam, and Tohidul Mazumder. "Mobile application and its global impact". In: *International Journal of Engineering & Technology* 10.6 (2010), pp. 72–78.

[2] Magdalena Stefańska and Tomasz Wanat. "Benefits from using mobile applications by Millennials–a gender and economic status comparative analysis". In: *The Proceedings of XVI International Marketing Trends Conference, Madrid, Paris-Venice Marketing Trends Association, available at: www. marketing-trends-congress. com/-papers*. 2017.

[3] Eudaldo Zamora-Intriago et al. "Use of mobile application as a means of communication of academic activities between parents and teachers". In: *2018 International Conference on Information Systems and Computer Science (INCISCOS)*. IEEE. 2018, pp. 221–226.

[4] José Luis Gutiérrez Rivas, Pedro Cano Olivares, and Javier Díaz Alonso. "Interactive techniques for entertainment applications using mobile devices". In: *Advances in Computational Intelligence: 12th International Work-Conference on Artificial Neural Networks, IWANN 2013, Puerto de la Cruz, Tenerife, Spain, June 12-14, 2013, Proceedings, Part II 12*. Springer. 2013, pp. 334–345.

[5] Cuauhtemoc Luna-Nevarez and Enda McGovern. "On the use of mobile apps in education: The impact of digital magazines on student learning". In: *Journal of Educational Technology Systems* 47.1 (2018), pp. 17–31.

[6] Athanasios S Drigas and Marios A Pappas. "A review of mobile learning applications for mathematics." In: *International Journal of Interactive Mobile Technologies* 9.3 (2015).

[7] Statista statistics. "Number of smartphone mobile network subscriptions worldwide from 2016 to 2022, with forecasts from 2023 to 2028". In: *Statista-Technology Telecommunications, https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide* (2023).

[8] Magdalena Stefańska and Tomasz Wanat. "Benefits from using mobile applications by Millennials–a gender and economic status comparative analysis". In: *The Proceedings of XVI International Marketing Trends Conference, Madrid, Paris-Venice Marketing Trends Association, available at: www. marketing-trends-congress. com/-papers*. 2017.

[9] Aijaz Ahmad Sheikh et al. "Smartphone: Android Vs IOS". In: *The SIJ Transactions on Computer Science Engineering & its Applications (CSEA)* 1.4 (2013), pp. 141–148.

[10] Nguyen Phuc Khang. "Zalo Mini App International University Union - Effective information connection solution for union members and young people". In: *School Level Science Conference "Promoting the Role of the Youth Union of Universities and Colleges in Improving Digital Capacity and Innovation for Students in Ho Chi Minh City"* (2023).

[11] Alessandro Biessek. *Flutter for Beginners: An introductory guide to building cross-platform mobile applications with Flutter and Dart 2*. Packt Publishing Ltd, 2019.

[12] William Danielsson. "React Native application development". In: *Linköpings universitet, Swedia* 10.4 (2016), p. 10.

[13] Kirthi Premadasa and Kavita Bhatia. "Real life applications in mathematics: What do students prefer?" In: *International Journal for the Scholarship of Teaching and Learning* 7.2 (2013), p. 20.

[14] Rene Vidal et al. "Mathematics of deep learning". In: *arXiv preprint arXiv:1712.04741* (2017).

[15] Raniah Zaheer and Humera Shaziya. "A study of the optimization algorithms in deep learning". In: *2019 third international conference on inventive systems and control (ICISC)*. IEEE. 2019, pp. 536–539.

[16] Khalid Al-Sulaiti, Osama Aldereai, and Imran Bashir Dar. "Application of business mathematics in finance, marketing, tourism and behavioural sciences: A mini review". In: *Applications of Mathematical Sciences* 1.2 (2022), pp. 57–66.

[17] Gerald C Hsu. "Using math-physical medicine to analyze metabolism and improve health conditions". In: *Diabetes Complications* 2.4 (2018), pp. 1–6.

[18] Adam Gamoran and Eileen C Hannigan. "Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school". In: *Educational Evaluation and Policy Analysis* 22.3 (2000), pp. 241–254.

[19] Shunji Mori, Hirobumi Nishida, and Hiromitsu Yamada. *Optical character recognition*. John Wiley & Sons, Inc., 1999.

[20] Shunji Mori, Hirobumi Nishida, and Hiromitsu Yamada. *Optical character recognition*. John Wiley & Sons, Inc., 1999.

[21] Wojciech Bieniecki, Szymon Grabowski, and Wojciech Rozenberg. "Image preprocessing for improving ocr accuracy". In: *2007 international conference on perspective technologies and methods in MEMS design*. IEEE. 2007, pp. 75–80.

[22] Gourab Nath et al. "Isolated OCR For Handwritten Forms: An Application in the Education Domain". In: (2022).

[23] Ishfaq Ahmad Bhat. "MATHEMATICS ANXIETY AMONG STUDENTS: AN OVERVIEW". In: (2020).

[24] K Abdul Gafoor and Abidha Kurukkan. "Why High School Students Feel Mathematics Difficult? An Exploration of Affective Beliefs." In: *Online Submission* (2015).

[25] Mohammed KA Kaabar. *A Friendly Introduction to Differential Equations*. Vol. 1. CreateSpace Independent Publishing Platform, 2015.

[26] Luis R Pino-Fan et al. "Analysis of the meanings of the antiderivative used by students of the first engineering courses". In: *International Journal of Science and Mathematics Education* 16 (2018), pp. 1091–1113.

[27] Simone Jablonski and Matthias Ludwig. "Teaching and Learning of Geometry—A Literature Review on Current Developments in Theory and Practice". In: *Education Sciences* 13.7 (2023), p. 682.

[28] James F Epperson. *An introduction to numerical methods and analysis*. John Wiley & Sons, 2021.

[29] Mathematical Programming Glossary. *The Nature of Mathematical Programming Archived*. Wayback Machine, 2014.

[30] Luc Pronzato, Henry P Wynn, and Anatoly A Zhigljavsky. "A generalized golden-section algorithm for line search". In: *IMA Journal of Mathematical Control and Information* 15.2 (1998), pp. 185–214.

[31] Rob Stevenson and Johannes Storn. "Interpolation operators for parabolic problems". In: *Numerische Mathematik* 155.1-2 (2023), pp. 211–238.

[32] Adi Ben-Israel. "A Newton-Raphson method for the solution of systems of equations". In: *Journal of Mathematical analysis and applications* 15.2 (1966), pp. 243–252.

[33] Elika Kurniadi, Zulkardi Zulkardi, and Ratu Ilma Indra Putri. "Learning ordinary differential equation at undergraduate level: A systematic learning review". In: *Al-Jabar: Jurnal Pendidikan Matematika* 13.1 (2022), pp. 23–31.

[34] BN Biswas et al. "A discussion on Euler method: A review". In: *Electronic Journal of Mathematical Analysis and Applications* 1.2 (2013), pp. 2090–2792.

[35] Geeta Arora, Varun Joshi, and Isa Sani Garki. "Developments in Runge–Kutta Method to Solve Ordinary Differential Equations". In: *Recent Advances in Mathematics for Engineering* (2020), pp. 193–202.

[36] AF Abdulhameed and QA Memon. "An improved Trapezoidal rule for numerical integration". In: *Journal of Physics: Conference Series*. Vol. 2090. 1. IOP Publishing. 2021, p. 012104.

[37] Eric W Weisstein. "Simpson's rule". In: *https://mathworld. wolfram. com/* (2003).

[38] Simon Saunders. "What is probability?" In: *Quo vadis quantum mechanics?* Springer, 2005, pp. 209–238.

[39] Charles Ezeagba. *Architecture Application Model View Controller (MVC) in Designing Information System of MSME Financial Report April 2017*. 2018.

[40] Ruiqi Yang and Eric Modesitt. "ViT2EEG: Leveraging Hybrid Pretrained Vision Transformers for EEG Data". In: *arXiv preprint arXiv:2308.00454* (2023).

[41] Kaiming He et al. "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[42] Anthony Gillioz et al. "Overview of the Transformer-based Models for NLP Tasks". In: *2020 15th Conference on Computer Science and Information Systems (FedCSIS)*. IEEE. 2020, pp. 179–183.

[43] Alexey Dosovitskiy et al. "An image is worth 16x16 words: Transformers for image recognition at scale. arXiv 2020". In: *arXiv preprint arXiv:2010.11929* (2010).

[44] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. "Layer normalization". In: *arXiv preprint arXiv:1607.06450* (2016).

[45] Khiem T. Do. "Handwritten_mathematical_expressions_recognition-". In: *GitHub, https://github.com/rockydtk/Handwritten_mathematical_expressions_recognition-* (2020).

[46] Lukas Bletcher. "LaTeX-OCR". In: *GitHub, https://github.com/lukas-blecher/LaTeX-OCR* (2021).

[47] Green Wood. "CoMER: Modeling Coverage for Transformer-based Handwritten Mathematical Expression Recognition". In: *GitHub, https://github.com/Green-Wood/CoMER* (2021).

[48] Anh Duy Le. "A Hybrid Vision Transformer Approach for Mathematical Expression Recognition". In: *International Conference on Digital Image Computing: Techniques and Applications (DICTA)* (2022).

[49] Guido Van Rossum et al. "Python Programming Language." In: *USENIX annual technical conference*. Vol. 41. 1. Santa Clara, CA. 2007, pp. 1–36.

[50] Damian Podareanu et al. "Best practice guide-deep learning". In: *Partnership for Advanced Computing in Europe (PRACE), Tech. Rep* 2 (2019).

[51] Rachel K Luu and Markus J Buehler. "BioinspiredLLM: Conversational Large Language Model for the Mechanics of Biological and Bio-Inspired Materials". In: *Advanced Science* (2023), p. 2306724.

[52] lucidrains. "x-transformer". In: *GitHub, https://github.com/lucidrains/x-transformers* (2023).

[53] Archan Ghosh. "Linear Equation Dataset". In: *Kaggle, https://www.kaggle.com/datasets/archa equation-images* (2021).

[54] Archan Ghosh. "Images of Randomly Generated Quadratic Equations". In: *Kaggle, https://www.kaggle.com/datasets/archanghosh/images-of-randomly-generated-quadratic-equations* (2021).

[55] CROHME. "CROHME Dataset". In: *Website, https://crohme2023.ltu-ai.dev/data-tools/* (2014, 2016).

Chapter 6

Appendix

6.1 External links

The final purpose is to approach the customers, who are mostly students, it is necessary for AiMA - a mathematics mobile application to be deployed for use. Concurrently, the first version of this app is available on the Play Store. AiMA can be accessed at here, or by scanning the following QR code

Due to the first version of this application, there absolutely needs to be more improvements in the future.
About the source of code, it is accessible on GitHub, for concerns, here is the link for the GitHub repository.

Chapter 7

Appendix

7.1 Publications and awards

This application has proven its performance and scientificity when it has been accepted for publication by conferences and a journal. Specifically, this topic has been accepted for posting at:

Journal

- **Nam Anh Dang Nguyen**, Binh Nguyen Le Nguyen, Duc Dang Khoi Nguyen, Le Duy Tan: "AiMA - An AI-Based Mobile System to Assist College Students with Math-Related Issues, Science and Technology Development Journal", Accepted (July 2023).

Conference

- **Nam Anh Dang Nguyen**, Binh Nguyen Le Nguyen, Le Duy Tan: "AiMA - An AI-Powered Mobile System for Supporting Undergraduate Dealing with Mathematical Problems, The 7th International Student Science Forum 2023 (ISSF 2023)". Accepted (September 2023).

- **Nam Anh Dang Nguyen**, Binh Nguyen Le Nguyen, Le Duy Tan: "AiMA - An AI-Based Mobile System to Assist College Students with Math-Related Issues, The Conference on The Digital Transformation Capacity for Vietnamese Youth 2023 – Vietnam". Accepted (May 2023).

Awards

- First Prize in The conference on the Digital Transformation Capacity for Vietnamese Youth 2023 with the article "AiMA - An AI-Based Mobile System to Assist College Students with Math-Related Issues, The Conference on The Digital Transformation Capacity for Vietnamese Youth 2023 – Vietnam"

- Top 3 best research articles at The 7th International Student Science Forum 2023 (ISSF 2023) with the article "AiMA - An AI-Based Mobile System to Assist College Students with Math-Related Issues, The Conference on The Digital Transformation Capacity for Vietnamese Youth 2023 – Vietnam"