



INFORMATICS
INSTITUTE OF
TECHNOLOGY

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

**MathQ -
Answer Generator for Sinhala Arithmetic Problems**

A dissertation by

Shenal Anthony

Supervised by

Mrs. Niwarthana Kariyabadhuge

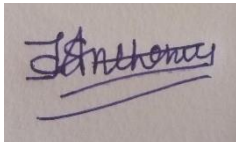
Submitted in partial fulfilment of the requirements for the BSc in Computer Science
degree at the University of Westminster.

Declaration

I hereby certify that this project study along with the objects which are related to the project is my own work and that it has not been submitted previously nor currently being submitted for any degree program.

Full Name : Congeneege Jude Shenal Anthony

Registration No : 2018383 | w1742306

A photograph of a handwritten signature in blue ink on a light-colored surface. The signature appears to be 'Shenal Anthony' written in a cursive style.

Signature :

Date : 25th of July 2022

Abstract

Natural Language Processing(NLP) has given the opportunity of letting the machines understand and perform various tasks over text and speech data involving different languages. The research is also based on Natural Language Processing. Generating answers for Sinhala Arithmetic Problems that contains multiplication and division. Sinhala language is only used in Sri Lanka as the first language. Sinhala arithmetic problems have a different way of expressing the characteristics of the problem comparing to other languages. Sinhala domain is a less explored area since it's a less resource domain. Answer generation for arithmetic problems that contains multiplication and division is the research area that would be discussed in this research.

As for the approach of the system MathQ has used Logistic regression for the question classification part. With that the related work the research will build the solution for generating answers for Sinhala arithmetic problems that contains multiplication and division. From identifying the keywords in the problem after identifying the operator then for the calculation of the answer using the NLP model. The NLP model will be trained and tested with the suitable data sets to get better accuracy and to get better performance.

The solution has an improvement comparing to the related work that has made in the same domain. The system has been tested over many times and have shown great results and have been optimized to get the better performance and higher accuracy. The model can optimize for further developments that may increase the scope of the system.

Keywords : Natural Language Processing, Machine Learning, Question Answering, Sinhala Mathematical Problems

Acknowledgement

So many difficulties, obstacles and hard times were there during the final project time period. First of all, I want to express my upmost gratitude to our God Almighty for providing me the much needed strength and positive mindset to overcome all the difficulties that I faced during the time. Also, I would like to express my heartfelt gratitude to my parents and my siblings, who have always been the most supporting and encouraging people in my life. I would like to express my heartfelt gratitude to my supervisor Mrs. Niwarthana Kariyabadhuge for her continues support during the project which I would not have been able to complete without her guidance. I would like to thank all the lectures, domain experts, colleagues and the Informatics Institute of Technology, for helping me to complete the project in many ways. Finally, and most importantly I want to thank me for believing in me and working hard to complete the project without quitting.

TABLE OF CONTENTS

Declaration.....	II
Abstract.....	III
Acknowledgement.....	IV
List of Figures.....	XI
List of Tables	XIII
List of Equations	XV
List of Abbreviations	XV
CHAPTER 1 : INTRODUCTION.....	1
1.1 Chapter Overview	1
1.2 Problem Domain	1
1.2.1 Sinhala Domain.....	1
1.2.2 Mathematical Domain.....	2
1.3 Problem Definition.....	2
1.3.1 Problem Statement.....	3
1.4 Research Motivation	3
1.5 Existing Work	4
1.5.1 Research Based Related Work.....	4
1.5.2 Domain Based Related Work.....	5
1.6 Research Gap	6
1.7 Contribution to the Body of Knowledge.....	6
1.7.1 Technological Contribution	6
1.7.2 Domain Contribution	7
1.8 Research Challenge.....	7
1.9 Research Questions	7
1.10 Research Aim.....	7

1.11 Research Objective	8
1.12 Project Scope	9
1.12.1 In-scope.....	9
1.12.2 Out-scope	9
1.12.3 Prototype Feature Diagram	10
1.13 Chapter Summary	10
CHAPTER 2 : LITERATURE REVIEW	11
2.1 Chapter Overview	11
2.2 Concept Map.....	11
2.3 Reviewing of the Problem Domain.....	11
2.3.1 Question Answering.....	11
2.3.2 Sinhala Domain.....	12
2.3.3 Mathematical Domain.....	13
2.4 Existing Work	14
2.5 Technologies Used in Existing Systems	20
2.5.1 Natural Language Processing	20
2.5.2 Conditional Random Field.....	22
2.5.3 K-Nearest Neighbor	22
2.5.4 Bidirectional Encoder Representations from Transformers(BERT)	22
2.5.5 Support Vector Machine	23
2.5.6 Naïve Bayes	23
2.6 Chapter Summary	24
CHAPTER 3 : METHDOLOGY.....	25
3.1 Chapter Overview	25
3.2 Research Methodology	25
3.3 Development Methodology	26

3.3.1 Design Methodology.....	26
3.3.2 Evaluation Methodology.....	26
3.4 Project Management Methodology.....	27
3.4.1 Deliverables	27
3.4.2 Gantt Chart.....	27
3.4.3 Resource Requirement	28
3.4.4 Risk and Mitigation.....	29
3.5 Chapter Summary	29
CHAPTER 4 : SOFTWARE REQUIREMENT SPECIFICATION	30
4.1 Chapter Overview	30
4.2 Rich Picture.....	30
4.3 Stakeholder Analysis	31
4.3.1 Stakeholder Onion Model	31
4.3.2 Stakeholder Viewpoints	31
4.4 Analysis of Requirement Elicitation Methodology	32
4.5 Discussion of Results.....	33
4.5.1 Findings from Literature Review Existing Systems	33
4.5.2 Questionnaire Findings	33
4.5.3 Findings from Brainstorming.....	41
4.5.4 Findings from Observation	41
4.6 Summary of Findings.....	42
4.7 Context Diagram.....	42
4.8 Use Case Diagram.....	43
4.9 Use Case Description.....	43
4.10 Requirement Specification.....	46
4.10.1 Functional Requirements	46

4.10.2 Non-Functional Requirements	47
4.11 Chapter Summary	47
CHAPTER 5 : SOCIAL, LEGAL, ETHICAL AND PROFESSIONAL ISSUES	48
5.1 Chapter Overview	48
5.2 Social, Legal, Ethical and Professional Issues	48
5.3 Chapter Summary	49
CHAPTER 6 : SYSTEM ARCHITECTURE & DESIGN	50
6.1 Chapter Overview	50
6.2 Design Goals	50
6.3 System Architecture Design	51
6.3.1 Layered Architecture	51
6.4 System Design	52
6.4.1 Design Paradigm	52
6.4.2 Data Flow Diagram	53
6.4.3 Activity Diagram	54
6.4.4 UI Designs	54
6.5 Chapter Summary	54
CHAPTER 7 : IMPLEMENTATION	55
7.1 Chapter Overview	55
7.2 Technology Selections	55
7.2.1 Technology Stack	55
7.2.2 Data Selection	56
7.2.3 Programming Language	56
7.2.4 Libraries	56
7.2.5 Selection of Development Framework	57
7.2.6 Selection of IDE's	58

7.2.7 Selection of Version Control	58
7.2.8 Summary of Technology Selection.....	58
7.3 Core Functionality	59
7.3.1 Data Preprocessing.....	59
7.3.2 Data Vectorization	61
7.3.3 Classification.....	61
7.3.4 Answer Generation	62
7.4 Implementation of API's.....	63
7.5 User Interfaces	64
7.6 Chapter Summary	66
CHAPTER 8 : TESTING	67
8.1 Chapter Overview	67
8.2 Objectives and Goals of Testing	67
8.3 Testing Criteria	67
8.4 Model Testing	68
8.4.1 Confusion Matrix	68
8.4.2 Accuracy, F1 Score, Precision, Recall.....	69
8.5 Benchmarking	69
8.6 Functional Testing	70
8.7 Module and Integration Testing.....	71
8.8 Non-Functional Testing	72
8.8.1 Performance Testing	72
8.8.2 Scalability	73
8.9 Limitations of the Testing Process.....	73
8.10 Chapter Summary	73
CHAPTER 9 : EVALUATION.....	74

9.1 Chapter Overview	74
9.2 Evaluation Methodology and Approach	74
9.3 Evaluation Criteria	74
9.4 Self-Evaluation	75
9.5 Selection of the Evaluators	76
9.6 Evaluation Result	77
9.6.1 Expert Opinion	77
9.6.2 Focus Group Testing	80
9.7 Limitations of Evaluation	80
9.8 Evaluation on Functional Requirements	81
9.9 Evaluation on Non-Functional Requirements	82
9.10 Chapter Summary	82
CHAPTER 10 : CONCLUSION.....	83
10.1 Chapter Overview	83
10.2 Achievements of Research Aims & Objectives	83
10.2.1 Achievement of the Research Aim	83
10.2.2 Achievement of Research Objectives	83
10.3 Utilization of Knowledge from the Course	84
10.4 Use of Existing Skills	84
10.5 Use of New Skills	85
10.6 Achievement of Learning Outcomes	85
10.7 Problems and Challenges Faced	86
10.8 Deviations	86
10.9 Limitations of the Research	87
10.10 Future Enhancements	87
10.11 Achievement of the Contribution to Body of Knowledge	88

10.12 Concluding Remarks.....	89
References.....	i
Appendix A – Concept Graph	v
Appendix B – Gantt Chart	vi
Appendix C – Activity Diagram	vii
Appendix D – UI Designs	viii
Appendix E – Requirement Gathering Questionnaire	ix
Appendix F – Evaluation Question Email	xii
Appendix G – Evaluation Responses of Expert Evaluators.....	xii

List of Figures

Figure 1.1 - Prototype Feature Diagram	10
Figure 2.1 - Overview of the BIO-Answer Finder.....	15
Figure 2.2 - Answer processing of Mahoshadha	17
Figure 2.3 - Sinhala Tag Set (Jayaweera and Dias, 2022)	19
Figure 2.4 - Identifying the hyperplane	23
Figure 4.1 - Rich Picture	30
Figure 4.2 - Stakeholder Onion Model	31
Figure 4.3 - Questionnaire Response 1	33
Figure 4.4 - Questionnaire Response 2	34
Figure 4.5 - Questionnaire Response 3	35
Figure 4.6 - Questionnaire Response 4	36
Figure 4.7 - Questionnaire Response 5	37
Figure 4.8 - Questionnaire Response 6	37
Figure 4.9 - Questionnaire Response 7	38
Figure 4.10 - Questionnaire Response 8	39

Figure 4.11 - Questionnaire Response 9	40
Figure 4.12 - Context Diagram	42
Figure 4.13 - Use Case Diagram.....	43
Figure 6.1 - Layered Architecture.....	51
Figure 6.2 - Level-1 Data Flow Diagram.....	53
Figure 6.3 - Level-2 Data Flow Diagram.....	54
Figure 7.1 - Technology Stack.....	55
Figure 7.2 - Label Dataset.....	59
Figure 7.3 : Normal Dataset.....	60
Figure 7.4 - Data preprocessing	60
Figure 7.5 - Assigning training and testing data	61
Figure 7.6 - Data vectorization	61
Figure 7.7 - Training the regression model.....	61
Figure 7.8 - Creating the pipeline	62
Figure 7.9 - Tokenization and POS tagging.....	62
Figure 7.10 - Keyword Extraction	62
Figure 7.11 - Predicting the question category	62
Figure 7.12 - Answer Calculation.....	63
Figure 7.13 - Project structure	63
Figure 7.14 - API routes.....	64
Figure 7.15 - Home page	65
Figure 7.16 - Generate Answer Page	65
Figure 7.17 - Answers Page	66
Figure 8.1 - Confusion Matrix	68
Figure 8.2 - Accuracy, F1 Score, Precision, Recall	69
Figure 8.3 - System Accuracy.....	69

Figure 8.4 - Performance testing.....	72
---------------------------------------	----

List of Tables

Table 1.1 - Research Based Related Work	5
Table 1.2 -Domain Based Related Work	6
Table 1.3 - Research Objective	9
Table 3.1 - Research Methodology	26
Table 3.2 - Deliverables	27
Table 3.3 - Risk and Mitigation	29
Table 4.1 - Stakeholder Viewpoints.....	32
Table 4.2-Findings from Literature Review Existing Systems.....	33
Table 4.3 - Questionnaire Findings 1	34
Table 4.4 - Questionnaire Findings 2.....	34
Table 4.5 - Questionnaire Findings 3.....	35
Table 4.6 - Questionnaire Findings 4.....	36
Table 4.7 - Questionnaire Findings 5.....	37
Table 4.8 - Questionnaire Findings 6.....	38
Table 4.9 - Questionnaire Findings 7.....	39
Table 4.10 - Questionnaire Findings 8.....	39
Table 4.11 - Questionnaire Findings 9.....	41
Table 4.12 -Findings from Brainstorming	41
Table 4.13 -Findings from Observation.....	41
Table 4.14 - Summary of Findings	42
Table 4.15 - Use Case Description UC-01.....	43
Table 4.16- Use Case Description UC-02.....	44
Table 4.17- Use Case Description UC-03.....	44

Table 4.18 - - Use Case Description UC-04	45
Table 4.19 - Use Case Description UC-05.....	45
Table 4.20 - Requirement Specification	46
Table 4.21 - Functional Requirements.....	47
Table 4.22 -Non-Functional Requirements.....	47
Table 5.1 - Social, Legal, Ethical and Professional Issues	49
Table 6.1- Design Goals	50
Table 7.1 - Libraries.....	57
Table 7.2 - Why use Flask	57
Table 7.3 - Selection of Version Control	58
Table 7.4 - Summary of Technology Selection	59
Table 8.1 - Benchmarking.....	70
Table 8.2 - Functional Testing	71
Table 8.3 - Module and Integration Testing	72
Table 9.1 - Evaluation Criteria.....	75
Table 9.2 - Self-Evaluation	75
Table 9.3 - Selection of the Evaluators	76
Table 9.4 - Project Context	77
Table 9.5 - Project Scope & Depth	78
Table 9.6 - Implementation of the Prototype	79
Table 9.7 - Context of the Thesis.....	80
Table 9.8 - Prototype Features Testing	80
Table 9.9 - Usability Testing	80
Table 9.10 - Evaluation on Functional Requirements	81
Table 9.11 - Evaluation on Non-Functional Requirements	82
Table 10.1 - Achievement of Research Objectives.....	84

Table 10.2 -Utilization of Knowledge from the Course84

Table 10.3 - Problems and Challenges Faced86

List of Equations

Equation 2.1 - CRF formula for text classification (Macherla, 2018)22

List of Abbreviations

Acronym	Description
CNN	Convolutional Neural Network
QA	Question and Answering
NLP	Natural Language Processing
BERT	Bidirectional Encoder Representations from Transformers
SVM	Support Vector Machine
SSADM	Structured System Analysis and Design Methodology
OOAD	Object-Oriented Analysis and Design
UI	User Interface
KNN	K-Nearest Neighbors
CRF	Conditional Random Fields
POS	Part of Speech Tagging

CHAPTER 1 : INTRODUCTION

1.1 Chapter Overview

The chapter would discuss the domain of the research, the identified problem in the research and the identified research gap. The contribution to the domain as well as for the technology would be discussed as what this research would be contributing to the body of knowledge. The different research challenges, research aim and the research objectives which will be discussed according to the learning outcomes.

1.2 Problem Domain

Answer generations are a research developed with the answer finder. Answer finder is a study area that has been developing throughout the years. Answer finder using corpus are the main research that has been developed to many languages. Also, the same research has been conducted using different technologies(Jayakody et al., 2016). Likewise answer generation for arithmetic problems are a developed research question from the research answer finder. The research can be extended to different languages since not all languages have the same sort of data when it comes to arithmetic problems. This research identify the need of documenting the different domains that would be addressing.

1.2.1 Sinhala Domain

Sinhala language is only used by Sri Lankans as the first language. And many use Sinhala as the second language to. Most of the schools in Sri Lanka have their education based on Sinhala language. Also, in Sri Lanka many of the industries are used to keep the language based documents handwritten rather than keep it in typed documents. Hence the software industry there are not many research done based on the Sinhala language. The reason for it would be the domain is a low resource domain comparing to other domains. To keep a record of work using Sinhala in a typed document is used less since there aren't many products that could do this other than mobile phones.

Very limited data are available through the internet on Sinhala language in the present. Another reason for it would be comparing with English language Sinhala language consists of sorts of

grammar involved that are hard to map out. So much research that are based on Sinhala language would have to create their own data set using different resources.

As per (Chathurika et al., 2016) there are not many Sinhala QA systems that are published on the internet for research. The Sinhala language domain opens up many research gaps in the future as well as now when its still developing as researchers might focus on the domain and conduct research in the future.

1.2.2 Mathematical Domain

Maths is a very well-known subject around the world. Since the ancient times Maths has been the key to development in technological aspects. Many countries has its own way of learning Maths and with different technics. Above all the basic operations that are used in Maths are Addition, Subtraction, Multiplication and Division. Those operations that are called arithmetic operations are the main key word a mathematical problem is created (Hosseini *et al.*, 2014)(Hosseini *et al.*, 2014).

Much research have been conducted in the past years under this domain in various industries. Maths is a common module/subject to every language in every country. Many arithmetic problems that are based on Maths comes under the operations of Addition, Subtraction, Multiplication and Division.

For different languages involving Mathematical domain based research have proven many ways to solve different problem statements throughout the past years. But few have been published involving Sinhala language. Many are done for the English language.

(Chathurika et al., 2016) is one of those research that have been done for the Sinhala domain. They have only conducted their research for the arithmetic problems that involved Addition and Subtraction. It opens up for future research to research more about the domain and contribute to the domain as well as for the software industry.

1.3 Problem Definition

Solving Sinhala Arithmetic Problems have a vast research area. Throughout the elementary schools to higher education in Sri Lanka Maths subject have different problems based on different concepts. This research would be mainly focusing on the arithmetic problems of

mathematics in Sinhala language. Out of that also multiplication and division arithmetic problems would be the main focus.

Many research that have been conducted on question answering does not work on solving mathematical problems(Chathurika et al., 2016). The research have their main aim of answering the question with the text paragraph where the answer is in are also provided to the system. So, for mathematical problems the question arise when there is known paragraph to be served to get the answer.

With that a different approach is needed to for the mathematical questions.

1.3.1 Problem Statement

How to solve a Sinhala Mathematical Arithmetic Question that is based on multiplication and division?

Sinhala mathematical problems have many ways that have been constructed, this research will be focused on the arithmetic operations of multiplication and division. The main problem statement will be how to solve Sinhala mathematical arithmetic questions of multiplication and division.

1.4 Research Motivation

The author was motivated on conducting a research related to Natural Language Processing(NLP). Authors supervisor advised to look into low resources domain because there are research gaps that are available on low resource domain. So, the author was reading research papers to find research gap when he came across a research paper on Question Answering for Sinhala language arithmetic problem using neural network. In that research paper in the future works section, it was clearly stated that they have worked on the arithmetic operations of addition and subtraction so the system can be created for arithmetic operation of multiplication and division. Since there was research gap available the author was able to find that research gap and conduct the research on that.

1.5 Existing Work

Many research has been conducted on Question Answering System for different domain and languages using different approaches and technologies throughout the years. In this section the author has discussed the related work the algorithm used in the research, the limitations and the improvements made in the research.

1.5.1 Research Based Related Work

Citation	Brief	Algorithm used	Limitations	Improvements
(Ozyurt, Bandrowski and Grethe, 2021)	Bio-Answer Finder: a system to find answers to questions from biomedical texts	BERT based reranking	Applicable only to English Language Domain	The only research that is done for the biomedical domain.
(Purwarianti, Tsuchiya and Nakagawa, 2007).	A machine learning approach for Indonesian question answering system.	SVM machine learning algorithm	Cross Language Question Answering using a similar approach with the Indonesian monolingual QA	Question Answering system can be easily adapted to other limited resource language
(Yin et al., 2016)	Simple Question Answering by Attentive Convolutional Neural Network. This work focuses on answering single-relation	Convolutional Neural Network(CNN)	Can only be used to their language	A simple and effective entity linker that brings higher coverage of ground truth entities.

	factoid questions over Freebase			
(Hosseini <i>et al.</i> , 2014).	Learning to Solve Arithmetic Word Problems with Verb Categorization	ARIS , a new method for solving arithmetic word problems	<p>1. Creates a new method name called Aris to solve arithmetic problems</p> <p>2. ARIS learns to categorize verbs with 81.2% accuracy and is able to solve 77.7% of the problems questions.</p>	Only developed for Addition and subtraction

Table 1.1 - Research Based Related Work

1.5.2 Domain Based Related Work

Citation	Brief	Algorithm used	Limitations	Improvements
(Jayakody <i>et al.</i> , 2016)	“Mahoshadha”, the Sinhala Tagged Corpus Based Question Answering System	K-Nearest Neighbor Algorithms	Using summarization and document clustering the response time can be improved	98% of accuracy and with a high efficiency of generating the response
(Chathurika <i>et al.</i> , 2016)	Solving Sinhala arithmetic question using neural network	Naive Bayes Classification Conditional Random Field Neural Network	The system only provides answers only for addition	The only Sinhala based arithmetic problem solving system with 76% accuracy.

			and subtraction. The system accuracy of 76%	
--	--	--	---	--

Table 1.2 -Domain Based Related Work

1.6 Research Gap

Related work that have been conduct on Sinhala arithmetic question solving (Chathurika et al., 2016) have clearly mentioned that the system can be enhanced for the arithmetic operation for multiplication and division. The research has been conducted on the addition and subtraction arithmetic operation only. The author has stated that the performance can be increased also since their system has only 76% of accuracy. The system was tested using a dataset of 100 since it's a low resource research. Since there is a clear performance gap this research would be addressing it and will propose a solution for it

1.7 Contribution to the Body of Knowledge

The main purpose of the research is to contribute to the body of knowledge. The contribution can be identified in two ways. The contribution to the technology and the contribution for the domain.

1.7.1 Technological Contribution

The main contribution in this area would be using the approach mentioned in the research paper(Chathurika et al., 2016) for the multiplication and division arithmetic operation and prove that the Convolutional Neural Network (CNN) approach can be used to archive the project aim. The approach that was used for this would be replaced with a new approach so that the performance of key word identification can be more accurate. The future researchers can get an idea on the research and could develop more on the research and could contribute more.

1.7.2 Domain Contribution

The contribution for the two domains Sinhala language domain and Mathematical domain where the research is conducted for help the domain to have more work on it as well as to help future researchers to conduct more research on the same domain. The Sinhala arithmetic problem solving involving multiplication and division operation would be the first time in these domains would be another contribution for the domain as well.

1.8 Research Challenge

The main challenge in this research would be the domain is a low resource domain. The Sinhala language is a low resource domain meant the data-set is hard to find and not many publications can be found for the same domain. The challenge of applying the research approach to achieve the research aim.

Creating a model for the implementation would be a challenging task since the author is not familiar with the technologies that will be used in the research. After the model implementation the training of the model and testing of it would be a bit of a challenge. The cause for it would be there should be enough data records for both of them. And during the model training the author would come up with the challenge of getting the right output without any errors that would also be a challenge during this research.

1.9 Research Questions

RQ1: How to use neural network to generate answers for Sinhala arithmetic questions of multiplication and division?

RQ2: How to overcome the lack of data resources in the domain?

RQ3 : What are the skills that would be improved or gained at the end of the research?

1.10 Research Aim

The aim of the research is to create a problem solving system for Sinhala arithmetic problems that consists of multiplication and division.

As per the research gap mentioned in the section 1.6 the main aim would be to apply the related research to build the system where the multiplication and division arithmetic operations based questions involved in Sinhala language to generate the answer.

1.11 Research Objective

Objective	Explanation	Learning Outcome
Problem Identification	Go through the related work and identify their research and implementation Getting the required knowledge to identify the problem Finding a suitable research gap to conduct the research.	LO1, LO2
Literature Review	How to use Neural Network to approach the project aim Analyze the related work and identify their approach in their research. The different technologies that can be used to approach the research aim How related researchers have improved the performance in their systems	LO2, LO3, LO4
Data Gathering and Analysis	Need to web scrape for online web sites to gather data Need to collect as many records of data as possible	LO4
Research Design	Create different designs to help the implementation	LO7

	Design working prototype before the implementation	
Implementation	<p>Work through the technologies and according to the designs start implementing prototypes for the system</p> <p>Through the prototype create the final system according to the requirements.</p> <p>Creating the final model to fulfill the research aim</p>	LO7
Testing and Evaluation	<p>Create automate testing to evaluate the written code</p> <p>Identify the bugs and solving them and enhancing the system along the way</p> <p>See to all the requirements are met and fully functional</p> <p>Improving the performance of the system</p>	LO7

Table 1.3 - Research Objective

1.12 Project Scope

When it comes to arithmetic problem solving it's a broad topic which can be extend to many domains. Out of all of it other than English language other languages can be identified as low resources languages in these kind of projects. So, it would be easy to narrow down the research as bellow discussed.

1.12.1 In-scope

- The research would be done for the Sinhala language.
- Only addition, subtraction, multiplicity and division mathematical operations would use in the research.

1.12.2 Out-scope

- Not applicable for other languages.

- Only simple question can be solved
- Not all mathematical problems can be solved.

1.12.3 Prototype Feature Diagram

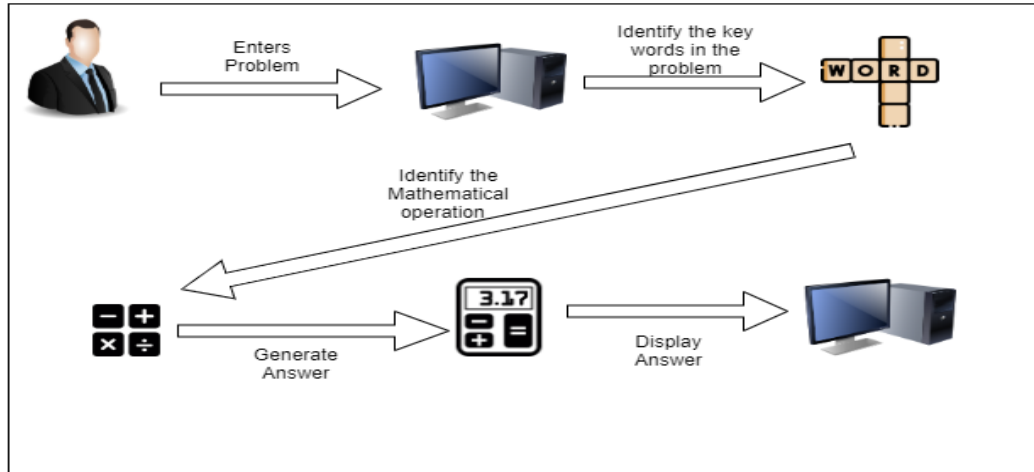


Figure 1.1 - Prototype Feature Diagram

1.13 Chapter Summary

The chapter began with explaining the domain of this research as the two domains of Sinhala language and Maths Domains were discussed. The problem definition was discussed along with the identified research gap then the contribution to the body of knowledge. The research challenge along with the research aim and research objectives were also discussed.

CHAPTER 2 : LITERATURE REVIEW

2.1 Chapter Overview

This chapter will cover the literature review of the research. Starting with the concept map including all the related works concepts in the diagram. The problem domain gives the needed knowledge about the research domain. Then after in the existing work section the different approaches that have been used along with the technologies used in those approaches are discussed.

2.2 Concept Map

Appendix A – Concept Graph

2.3 Reviewing of the Problem Domain

Three domains would be discussed in this section that includes the question answering domain, the Sinhala domain and the Mathematical domain.

2.3.1 Question Answering

The purpose of Question Answering(QA) systems is to get correct responses asked by the user using natural language. (Dwivedi and Singh, 2013). Many of the information retravels such as search engines allow users to ask question and then provide information regarding to the question. Most of the time the user should have to examine these information one by one in order to find the correct answer for the question. On the other hand, question answering systems gives the user the opportunity to ask the question and then get the answer for that particular question in very brief way. The time that would be spend overlooking into many resources would be decreased as the system would instantly give you the response.

Question answering systems can be divided into many types according to the question that is been asked from the system. The main types of question answering systems are ‘Open domain question answering systems’ and ‘Closed domain question answering systems’. (Reddy and Madhavi, 2017). In open domain QA systems, the user can ask any question regardless which

domain the question is in, as the system is not restricted to any domain. The world wide web is the biggest QA system that can answer any question that can be in any language or any domain such as health, politics, sports etc. Since the open domain QA systems mainly rely on open source resources such as Wikipedia the quality of the answer is not that high.

In close domain QA systems, the system is limited to a specific domain which the system is defined. In close domain QA systems, the user can only ask domain specific questions and can answer limited number of questions. (Reddy and Madhavi, 2017). Compared to the open domain, the close domain QA systems mainly rely on datasets or unstructured/ semi-structured data of the domain. So, the answers of these systems are accurate and the quality is very high compared to the open domain QA systems.

For this research the author is focused on close domain QA system. As the research scope is based on two domains. Sinhala domain would be the domain based on the language. As arithmetic problems are in the Mathematical domain, the questions will include both the Sinhala and Mathematical domains.

2.3.2 Sinhala Domain

Sinhala is the main language used by Sri Lankans, and the language is unique to the Sri Lankans as the language is not used by any other country. Comparing to the English language, Sinhala is more complicated language when it comes to grammar. In English alphabet there are 26 words while in Sinhala alphabet there are 58 letters and 38 words are used regularly. The Sinhala language is syllabic language where each consonant has its own vowel. (5 Fascinating Facts about the Sinhala Language - ALTA Language Services, 2019). The Sinhala alphabet is said to be descended from the Brahimi script which has long history in the world.

Mainly like in any other language Sinhala also follows the subject-object-verb word order in a sentence. The order is likely to change when applying the correct words in a sentence sometimes but most of the time the order is same as the English language. Moreover, Sinhala is a diglossic language where the everyday communication is different to the written one in literature, grammar. These two versions have their differences in pronunciation and in vocabulary as well. Native speakers are able to understand both the version and more often used version these days is the verbal one as it is easy and not bound to any grammar and can speak the language howsoever. The written version is bound to follow the grammatical rules

with the proper structure. As in this research the author is following the same set of rules with the creation of the questions.

The complexity of the language is there are many words that can be used to elaborate the meaning of a sentence in the same way. As reordering words in the sentence and giving out the same meaning is one specialty in Sinhala language. To demonstrate this look at the below two sentences,

- මා ලඟ පොත් 2ක් ඇත.
- පොත් 2ක් මා ලඟ ඇත.

Both of these two sentences give out the same meaning that is 'I have 2 books'. The difference of these two sentences are the order of which the words are connected. In English language the sentence 'I have 2 books', cannot change the order of the words and give out the same meaning and it is against the grammar rules but in Sinhala language as shown in the above two sentences the words are reordered and give out the same meaning and the sentences are following the necessary grammar rules as well. So, the complexity to show in this is someone who is learning the language would think out of the two sentences one is correct and the other one is not, as in most of the languages something like this is not possible. The same could be said for natural language processing, as the systems should be able to identify as these two sentences are the same.

2.3.3 Mathematical Domain

Mathematics is used by everyone in the world in different ways. Early in 2000BC it was said to be the start of Mathematics as the counting was introduce but it is also said that counting could not be named as mathematics when counting was kept as a record on something then it was identified as mathematics. (O'Connor and Robertson, 1997). Introducing the numbers made huge impact in humankind as it was developed time to time and this helped to archive a lot of great findings. As number problems like Pythagoras theorem were studied during the 1700BC period and then it made way to lot of linear equations problems were studied to solve problems that involve numbers.

With the mathematics were developing many countries have developed their own studies involving mathematics in their own language. Hence we get to see mathematical problems in different languages which we can see today. Although the language differentiate mathematical

problems the concepts for all those problems remains the same. Many of these mathematical problems consist basic mathematical operations. Which are ‘Addition’ , ‘Subtraction’, ‘Multiplication’ and ‘Division’. Even though most of problems does not have any of those basic operations in it the concept of which the problem is design falls into one of these operations.

Mathematical problems may have different ways of describing the problem as in a problem can be written in a paragraph or a sentence or it could be a straight equation. However, the problem appear the solution would always have numbers or letters or both in the end.

- සුමුදුගේ දැන් වයස අවුරුදු 9කි. ඇගේ අක්කාගේ දැන් වයස සුමුදුගේ දැන් වයස මෙන් 2 ගුණයකි. අක්කාගේ දැන් වයස කීය ද?
- 60m දිග ඉටි රෙද්දක් මේසවලට දැමීම සඳහා සමාන දිගින් යුතු කැබලි 2බැගින් කැබලි කොපමණ ප්‍රමාණයක් කැපිය හැකි ද?

As for the research the author is focus on sentence wise Sinhala mathematical problems which consist of the basic operations of ‘multiplication’ and ‘division’. The structure would contain the problem in Sinhala and numbers. The above given examples are some of the mathematical problems that would be used in this research.

2.4 Existing Work

Bert based Bio-Answer Finder

Bio-answer Finder (Ozyurt, Bandrowski and Grethe, 2021) is a research that was conducted to find aid bio medical researchers in order to answer their research questions. The research was a closed domain question answering system as the system is limited to the domain of bio medical. The preprocessing of the questions are done using the model called GloVE (Pennington, Socher and Manning, 2014). GloVE is a word embedding vector space model where it combines the global matrix factorization and local based word representation approaches.

The author has identified 3 phases in the process in order to find the answer, The first the question processing phrase, in this phrase a full analysis of the question is done to determine what the question is about. In the question processing phrase has the followings,

- Question Parsing – the questions are tokenized and then using part-of-speech(POS) tagging, lemmatization, constituent, and dependency parsing are done using the Stanford CoreNLP library.
- Question Type Detection – finding out the type of question that is there in order to find out the best way to answer it.
- Detection of the Focus of a Question – this stage would be to identify the necessary keywords to retrieve relevant document in order to find the answer

An overview of the system is given below to understand the system process in a light way.

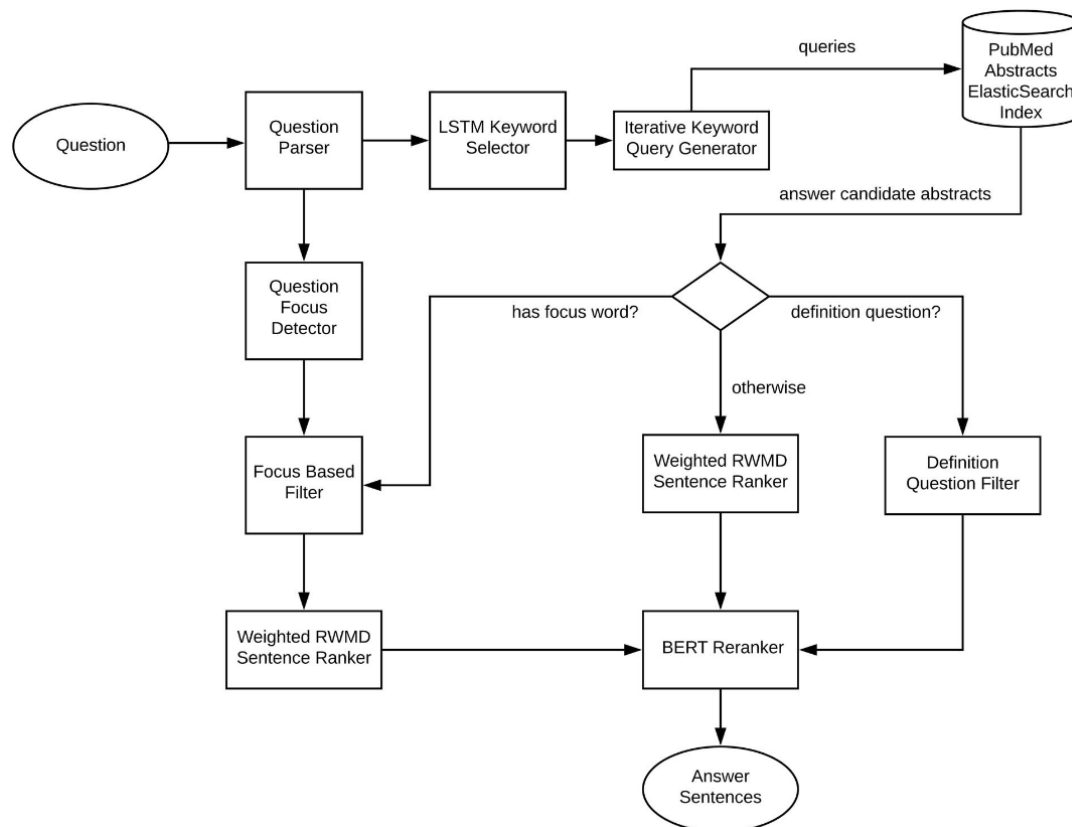


Figure 2.1 - Overview of the BIO-Answer Finder

The second phrase includes the document processing where the keywords are used to retrieve the relevant documents needed to find out the answer. The document retravel is guided by iterating Long-Shor-Term-Memory (LSTM) network.

The third and final phrase of the system is the answer processing phase, the supervised answer ranking is done using the BERT model. The model is used to weight the retrieved documents according to the question type identified during the question processing phase. After the fine-tuning is completed the abstract(the paragraph which is entered with the question) is indexed and the answer is given.

Support Vector Machine Based Indonesian Question Answer System

A machine learning approach for Indonesian QA System(Purwarianti, Tsuchiya and Nakagawa, 2007) is a close domain QA system which is only for the Indonesian domain. The approach the author has used is the Support Vector Machine widely known as SVM. The system is a couples based question answering system where both the question and the paragraph that contains the answer needs to be given to the system. Like all the other QA systems this also have the 3 main phases of ‘question classifier’, ‘passage retrieval’ and ‘answer finder’.

During the question classifier stage the system reads the question and using POS tagging the words are labeled and the main keywords are identified such as the noun, adjectives and verbs. After the tagging is completed the question is identified and categorized into the six expected answer types which are ‘location, date, person, organization, name or quantity’. The identification process is done using both ‘WordNet Distance’ and ‘Bi-gram Frequency’.

In the passage retriever stage, the relevant document is retrieved as per the identified keywords. The phase selects the most suited passage from the retrieved document. The answer finder stage is done using the SVM model as the model classify each word in the corpus according to the questions expected answer type and index the corpus. The answer would be presented after the model has done identifying the answer from the corpus.

K-Nearest Neighbor(k-NN) based Sinhala Question Answering System

“Mahoshadha” is a Sinhala tagged corpus based QA system (Jayakody et al., 2016). The research aim was to retrieve precise information from a large Sinhala tagged corpus. The research is closed domain worked on the Sinhala language domain. The research introduce a novel approach for Sinhala QA system. The approach mainly have the two stages of

summarizing the tagged corpus and using the summarization find the answer for the question. k-Nearest Neighbor(k-NN) Algorithms is used in order to cluster the summarized corpuses and SVM is used to automate both summarization and question understanding processing.

The system contains 4 components which are given below,

- Document summarizing – using SVM the documents are summarized in order get the categorizing of those documents done properly. Using text mining and recognizing text patterns the classifier summarize each of the documents.
- Document categorizing – the categorization is done using K-NN algorithm. Which the system is trained to categorize each document according to the content of the document(history, science, maths etc.)
- Question processing – the process is further divided into another two components,
 - Question analyze : POS tagging and language model(inform about the language to the POS tagger)
 - Identify answer type : Identifying the question(person type, location type, date type etc.)
- Answer processing – after the question type is identified the system would go through each paragraph looking for that type and along with the keyword search the answer would be extracted from that document. The process is given below for more understanding.

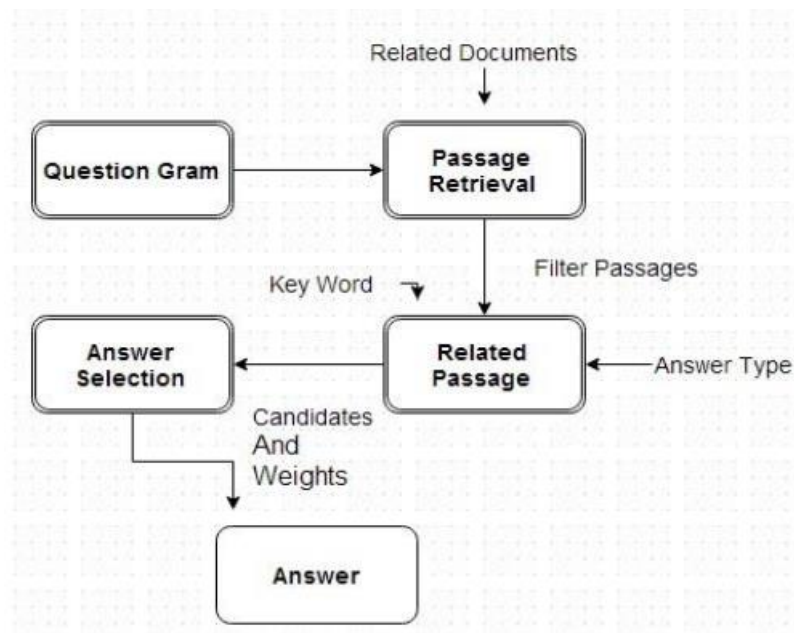


Figure 2.2 - Answer processing of Mahoshadha

The POS tagging in the system is done through the use of Hidden Markov Model this helps the system tag Sinhala texts more accurately.

Conditional Random Fields based Sinhala Arithmetic Problem Solving System

Solving Sinhala Language Arithmetic problems using Neural Network (Chathurika et al., 2016) is a research that is done one generating answers for arithmetic problems that consists of 'addition' and 'subtraction'. The approach has three phases 'keyword identification', 'question identification' and 'mathematical operation identification'. To identify the keywords in the question Naïve Bayes Classification has been used. To identify the question and the mathematical operation Conditional Random Fields has been used. One vs all classification is done using neural network that has built up equation to solve the problem.

The keyword identification process starts off with the use of POS tagging. This will help the system to identify each and every keyword in the sentence as the words are tagged as 'nouns', 'adjective', 'verb' etc.

According to (Jayaweera and Dias, 2022) below diagram shows the Sinhala tag set, the POS tagging would tag each word according to the below tag set.

	Tag	Description
1	NNR	Common Noun Root
2	NNM	Common Noun Masculine
3	NNF	Common Noun Feminine
4	NNN	Common Noun Neuter
5	NNPA	Proper Noun Animate
6	NNPI	Proper Noun Inanimate
7	PRPM	Pronoun Masculine
8	PRPF	Pronoun Feminine
9	PRPN	Pronoun Neuter
10	PRPC	Pronoun Common
11	QFNUM	Number Quantifier
12	DET	Determiner
13	JJ	Adjective
14	RB	Adverb
15	RP	Particle
16	VFM	Verb Finite Main
17	VNF	Verb Non Finite
18	VP	Verb Participle
19	VNN	Verbal Non Finite Noun
20	POST	Postpositions
21	CC	Conjunctions
22	NVB	Noun in Kriya Mula
23	JVB	Adjective in Kriya Mula
24	UH	Interjection
25	FRW	Foreign Word
26	SYM	Not Classified

Figure 2.3 - Sinhala Tag Set (Jayaweera and Dias, 2022)

The system has been trained using Naïve Bayes classifier to identify each and every word to be identified and tag them correctly. The next phase of question identification process is done using the Conditional Random field(CRF) tool. The tool helps to weight for each function and then after converting it into a probabilistic value in-between 0 and 1 which will help the system to identify if the question is asking for an addition or a subtraction. Along with the question identification the mathematical operation is also identified using the CRF tool. The numeric values in the questions would be extracted then after to do the mathematical operation on them to find out the answer.

Freebase Question Answering Using Relation Extraction and Textual Support

The research (Xu, Reddy, Feng and Zhao, 2016) was focused on creating a QA system on freebase via relation extraction and textual support. The system mainly consists of two components which are ‘inference on Freebase’ and ‘further inference on Wikipedia’.

- Inference on Freebase :-

During this phase the ‘entity linking’ and ‘relation extraction’ on the questions. During the entity linking stage the system is using a hand build sequences of POS to identify all the categories inside the question. For the entity linking the system is using a tool called ‘S-MART’ to retrieve the top entities from Freebase. After the retravel a ranking is conduct using a statistical model.

Next is the relation extraction where the system tries to identify the relation between the answer and the entities in the question. For this the system uses a Multi-Channel Convolutional Neural Network (MCCNN) which could exploit both syntactic and sentential information for relation extraction.

- Inference on Wikipedia

During this phase the answer selection is done according to the entities. For this the system uses Freebase API to convert Freebase entity to Wikipedia page. Using a refinement model the system search for the sentence that contains the answer entity. The model is treated as binary classification and LIBSVM (a library for SVM) is used for learning the weight the classification.

2.5 Technologies Used in Existing Systems

In this section the technologies and the algorithms used in the related work would be reviewed.

2.5.1 Natural Language Processing

Natural language processing(NLP) is a component of artificial intelligence which know for the computer’s ability to understand the human language that maybe in text or speech. (Lutkevich, Writer and Burns, 2022). NLP allow the computers to take human language by written or by speech and then convert it to a way that the computer can identify. NLP is mostly used for text

classification, text extraction, machine translation, natural language generations. In NLP there are two main phases in which the process flows. Which are 'data preprocessing' and 'algorithm development'.

Data Preprocessing

During the data preprocessing stage, the cleaning of the data is done before the machine could analyze it. The data would be text phrases where the preprocessing stage highlights certain features that is in the text so that the algorithms could work on it later on. The data preprocessing can be done in many different ways the list of ways are given below,

- **Tokenization** – The sentence is broken down into smaller units(word by word)
- **Stop word removal** – Common words that might be included in the sentence is removed so only the words with more information in the sentence remains.
- **Lemmatization-** The grouping of different forms of the same word can be done.
- **Stemming-** Allows the lowers inflection in words to their root forms
- **POS-** The tokenized words are tagged according to their own grammar phrases such as nouns, verbs, adjective, etc.
- **Word segmentation-** Separates the continuous text.

Algorithm Development

After the data preprocessing is completed the algorithm process the data according to the purpose of the system. Using NLP there are many ways to build algorithm to analyze the data, mainly there are two ways of using it

- **Ruled based systems-** The approach uses linguistic rules to process the data for the analysis.
- **Machine learning based systems-** The approach contains statistical methods to perform tasks based on training data. The methods are adjust as more training data is fed to the model.

2.5.2 Conditional Random Field

CRF is a model that is trained to do the task specific prediction like named entity recognition.(Chavan, 2019). CRF holds the ability to model the sequential data that can be used in NLP and many areas. Named entity recognition is a famous application in CRF where the model predicts the sequence in which the data is dependent on each other. The formula that is used in CRF text classification is given below, ‘Y’ is the hidden state and ‘X’ is the observed variable.

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \prod_{t=1}^T \exp \left\{ \sum_{k=1}^K \theta_k f_k(y_t, y_{t-1}, \mathbf{x}_t) \right\}$$

Equation 2.1 - CRF formula for text classification (Macherla, 2018)

CRF is an undirected graphical model where the nodes can be divided into two sets ‘x’ and ‘y’, then only the conditional distribution ‘p(y | x)’ is modeled. Higher-order, semi-Markov and Latent- dynamic are some CRF variants which are available.

2.5.3 K-Nearest Neighbor

K-Nearest Neighbor(KNN) is a supervised machine learning data classification method that can be trained to identify certain data pattern and group them or categorized it accordingly. KNN is mainly used to solve classification and regression problems. During the training phases the KNN model intend to store the data but only performs the task when a query is performed on the dataset. KNN is easy to develop and understand compared to other classifications on the other hand as the scale of the data is increasing the there is a decrement in the performance(slowing down). KNN use a voting mechanism to determine the class of an entity. According to the value of ‘K’ the class of a data point is defined and will use the nearest neighbors to it. (Joby, 2021).

2.5.4 Bidirectional Encoder Representations from Transformers(BERT)

Bidirectional Encoder Representations from Transformers(BERT) is based on transformers a deep learning model where every output element is connected to every input element. Based

on the connection the weighting are dynamically calculated. BERT is open source framework for NLP. The purpose of Bert is to let computers understand human language which are in text format with the use of surrounding texts. BERT is a pretrained model using text from Wikipedia and most likely suited to perform queries on question and answer datasets. (Lutkevich, 2020)

2.5.5 Support Vector Machine

Support vector machine(SVM) is machine learning algorithm that performs supervised learning for classification or regression. The main purpose of SVM is to find hyperplane in an N-dimensional space that clearly classify data point. In order to separate two classes of data points SVM would look for the maximum distance between data points of both classes. Compared to other classifications SVM can give the system a high accuracy with low consumption for and scale of data. (Gandhi, Support Vector Machine 2018)

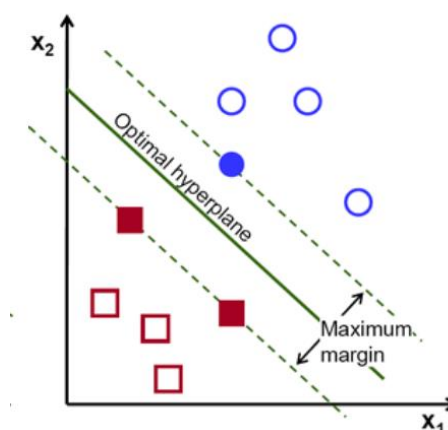


Figure 2.4 - Identifying the hyperplane

2.5.6 Naïve Bayes

Naïve Bayes is classification technique based on Bayes theorem with the assumption of a particular feature is unrelated to the presence of any other one. So, the classification falls under the probabilistic machine learning model. Naïve bayes is mostly used for sentimental analysis as it is easy to implement but the downfall of it that the requirement of predictors to be

independent. Mainly there are three type of Naïve bayes classifiers, Multinomial Naive Bayes, Bernoulli Naive Bayes, Gaussian Naive Bayes. (Gandhi, Naive Bayes Classifier 2018)

2.6 Chapter Summary

In this chapter the concept map showed the different concepts used in similar work and what concept will be used in this research. The domain section elaborate the knowledge about the research domain using previous work. Reviewing previous work was done under the existing work section where the different related work and their approaches along with the different technologies they have used were also discussed.

CHAPTER 3 : METHDOLOGY

3.1 Chapter Overview

The methodology chapter contains the different methods to overcome different stages of the research and comparing it with each to get the maximum effort and maximum outcome to make this research a success.

3.2 Research Methodology

Research Philosophy	Positivism will be used here since its important to conduct interviews with domain experts. Reviews of the domain experts are valuable in this research since it would be a good evaluation for the system.
Research Approach	Deductive Reasoning would be used here since it will do testing and evaluate the system and evaluate the accuracy. After the implementation the system would be tested to find the accuracy and evaluate the test results after work on to optimize the system for better performance.
Research Strategy	Not only limiting to the domain but expanding the requirement gathering to similar domains such as the English language domain. Since the research is based on low resource not many domain based research would be found so its better to refer to other domains research papers.
Research Choice	The choices are not to limit to one method of approach but try various methods and get the best one out of it.
Time zone	Cross-sectional time zone will be used since data collected

	beforehand. Before the prototyping the data sets would be collected and modified for the research purpose.
Techniques and Procedures	Table formats is suitable for techniques as its easy to write data and read. It is known that a reader can understand the content if it is in a table format. Since a tabular format is a structured way.

Table 3.1 - Research Methodology

3.3 Development Methodology

There are many methodologies that we could use to develop a project. For a research project the most suitable methodology would be prototyping methodology. The methodology would be chosen because during the implementation stage the prototype would be implemented and after testing the prototype and would work on the prototype until the wanted outcome is archived and a final system would be developed by the end of it.

3.3.1 Design Methodology

Since Structured System Analysis and Design Methodology(SSADM) is well suited for projects with stable requirements that won't change in time, the author would use SSADM as the design methodology for the research. The research have a solid research aim and the requirement would not change over the time. So, the suitable design methodology for a research would be SSADM.

3.3.2 Evaluation Methodology

Many evaluation metrics are available for research projects that are based on Neural Network. Taking consideration of related work and their evaluation methodology the author has chosen Classification Accuracy as the Evaluation Methodology. Classification accuracy methodology is vastly used methodology to calculate the accuracy of the system. Since it's easy to use the more test data you have the higher accuracy and less false sense of the accuracy you would get.

3.4 Project Management Methodology

Prototyping methodology would be used here as well. Since the author would have to design prototype and test it and repeat the sequence until the prototype is perfected and the accuracy is high in order for the product to get higher performance.

3.4.1 Deliverables

Deliverable	Date
Draft Project Proposal	23 rd September 2021
Literature Review Draft	21 st October 2021
PID Final Submission	4 th November 2021
SRS Draft	25 th November 2021
Interim Progress Report (IPR)	27 th January 2022
PSPD	3 rd March 2022
Test and Evaluation Report	17 th March 2022
Draft Project Reports	31 st March 2022
Final Project Report	5 th May 2022

Table 3.2 - Deliverables

3.4.2 Gantt Chart

Appendix B – Gantt Chart

3.4.3 Resource Requirement

Below are the requirements that are needed to complete the project.

3.4.3.1 Hardware Requirement

- Minimum ram should be 8 GB
- Minimum processor should be Core i5
- Minimum memory to run the system would be 2GB

3.4.3.2 Software Requirement

- The programming language would be python
- Draw.io to create diagrams, designs for the system
- Google docs Would be used to handle the documentation
- Git Hub would be used to handle the version controls
- Zotero for citation for the documentation.

3.4.3.3 Skill Requirement

- Best practice with NLP and using python programming and Web development
- Research writing skills for the reports

3.4.3.4 Data requirements

- Web scrape data set from the internet or create dataset from reliable resources.
- Create own dataset using the grade 5 mathematics textbook

3.4.4 Risk and Mitigation

Risk	Severity	Frequency	Mitigation
Lack of knowledge in key areas in the research	High	High	Read more research papers and do online courses to help gain the knowledge needed
Loss of recent work that has been done for the research	High	High	Always keep back in Cloud storage and always keep updating the Git Hub repository.
Lack of time to complete the research	Moderate	Moderate	Always be aware of the deadline ahead of time and work according to the Agile methodology to keep the workflow in a good way.

Table 3.3 - Risk and Mitigation

3.5 Chapter Summary

The chapter discussed the different methodologies in the research. The Research methodology, development methodology and in the project management methodology the project deliverables, the gantt chart and the risk and mitigations were identified and discussed.

CHAPTER 4 : SOFTWARE REQUIREMENT SPECIFICATION

4.1 Chapter Overview

This chapter contains on how system requirements are gathered and depth discussion of each requirement in the system. Starting with the rich picture diagram to get the flow of the system after the analysis of stakeholders of the system using the onion model. Then the analysis of requirement elicitation methodology is discussed. After the questionnaire findings and the summary of the findings are discussed. The context diagram along with the use case diagram and use case description are given to identify and analyze the system requirements. Finally, the functional and non-functional requirements are identified and discussed.

4.2 Rich Picture

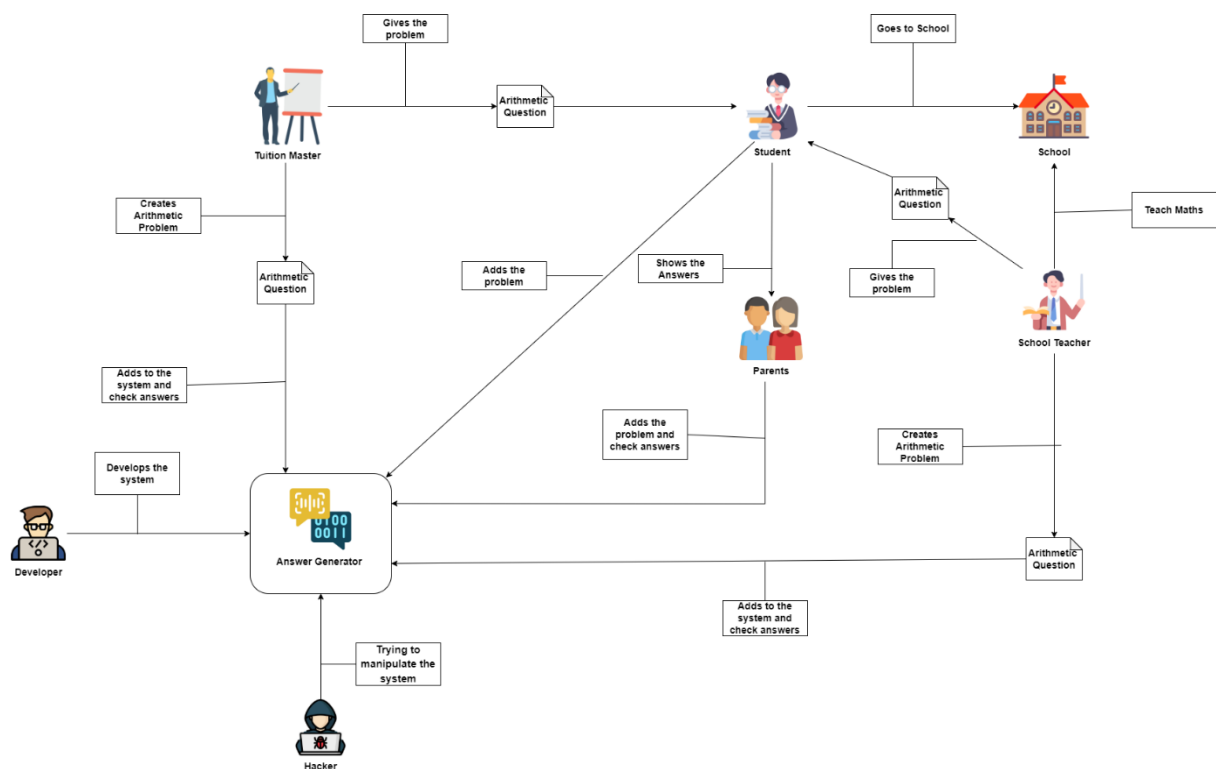


Figure 4.1 - Rich Picture

4.3 Stakeholder Analysis

Identifying the stakeholders and discussing their involvement to the system.

4.3.1 Stakeholder Onion Model

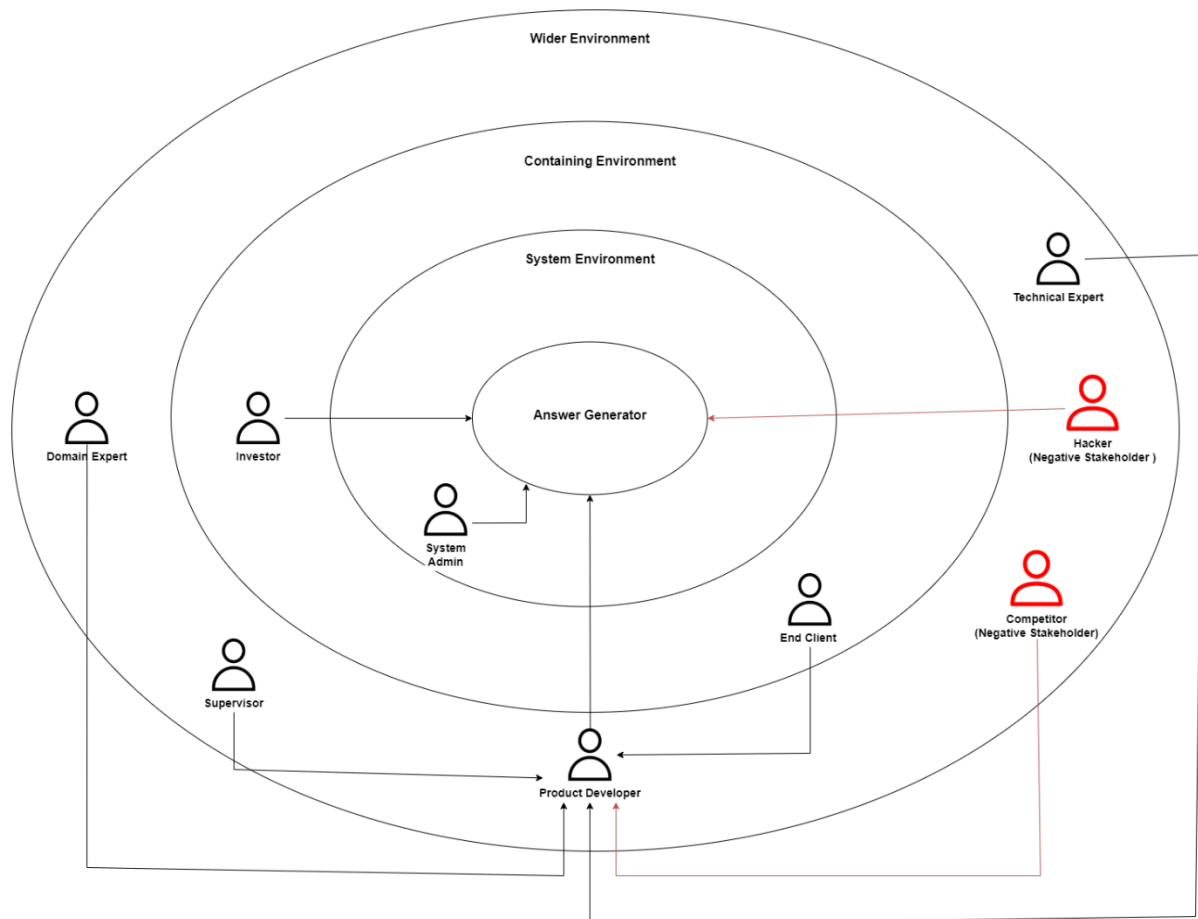


Figure 4.2 - Stakeholder Onion Model

4.3.2 Stakeholder Viewpoints

Stakeholder	Role	Descriptions
Product Developer	Functional Beneficiary	The primary actor and the stakeholder who develops the system.
End Client	Functional Beneficiary	Uses the final product.
Supervisor	Functional Beneficiary	Gives the necessary supervision to the developer for the development.

Domain Expert	Expert Stakeholder	Provide the recommended knowledge of the domains.
Technical Expert	Expert Stakeholder	Provide the knowledge on how to improve the system get the maximum output.
Investor	Financial Beneficiary	Provides the financial support to market the product.
System Admin	Operational Beneficiary	Keeps a track of the system after been deployed in different platforms.
Hacker	Negative Stakeholder	Try to manipulate the system and modify it content make the system work false data or false information.
Competitor	Negative Stakeholder	Try to create a system which is similar to the final product and try to degrade the system the author created.

Table 4.1 - Stakeholder Viewpoints

4.4 Analysis of Requirement Elicitation Methodology

Analysis of requirement elicitation is way of collecting requirements in the research. Gathering information for the research is very important since it will be very good way of looking into the research in another way. Conducting surveys, brainstorming, conducting interviews are some of the ways in which we can gather requirements/ information. In this research the author has gone with conducting surveys and brainstorming as a part of the requirement gathering for the research. The reason for this is surveys are very convenient way of asking the audience their opinion of the research as well as the system , not having to conduct interviews was with the amount of time that was spend in the research no time was left to conduct interviews for requirement gathering.

4.5 Discussion of Results

Discussing the results of different requirements gathering is important as it would help to develop the system to the up best way as possible.

4.5.1 Findings from Literature Review Existing Systems

Findings	Citations
Accuracy can be increased while optimizing the system as well as using a large set of data.	(Chathurika et al., 2016)
Created a new approach called 'ARIS' which solves English arithmetic questions with an accuracy of 77.7%	(Hosseini <i>et al.</i> , 2014).
Question answering system for Bio-medical text using BERT only used for English Language.	(Ozyurt, Bandrowski and Grethe, 2021)
The response time of the system can be increased using summarization and document clustering.	(Jayakody et al., 2016)

Table 4.2-Findings from Literature Review Existing Systems

4.5.2 Questionnaire Findings

Select your age group

150 responses

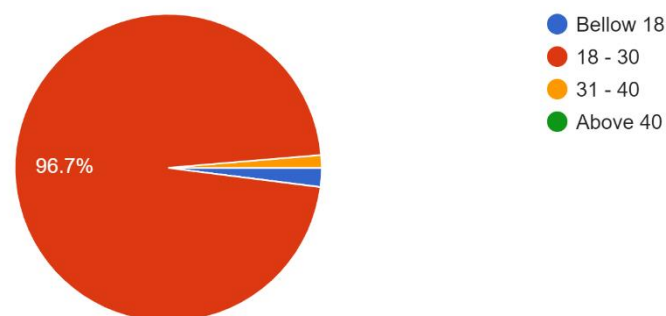


Figure 4.3 - Questionnaire Response 1

Question	Select your age Group
----------	-----------------------

Aim of the Question	To identify the responder's age
Observations	96.7% of responder's are in the age group of 18-30. Only 2% are in the age group of bellow 18 and 1.3% of them are in the age group of 31-40. No one has filled the survey that are in the age group of above 40.
Conclusion	Many of the responses have experience as they are above the age 18 and have the school knowledge and sure to have come across with the subject of the survey in the past.

Table 4.3 - Questionnaire Findings 1

Are you familiar with Sinhala Arithmetic Problems?

150 responses

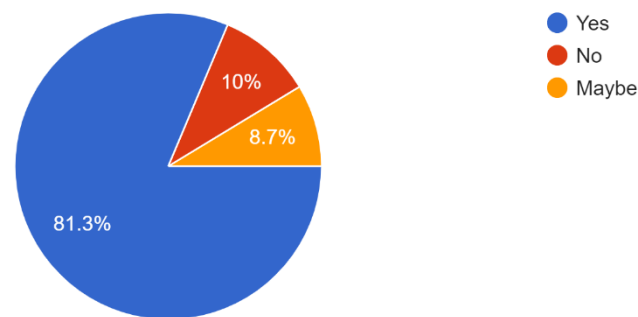


Figure 4.4 - Questionnaire Response 2

Question	Are you familiar with Sinhala Arithmetic Problems?
Aim of the Question	To identify if the responder has an idea of the subject.
Observations	81.3% have answered as Yes to the question while 10% have answered No and 8.7% have answered Maybe.
Conclusion	Many responders are familiar with the term Sinhala Arithmetic Problems and can tell that they have a full idea of the question that are in the survey.

Table 4.4 - Questionnaire Findings 2

On a scale of 1 to 5 (where 1 been the lowest and 5 been the highest) how would you scale your skills in solving Sinhala Arithmetic Problems?

150 responses

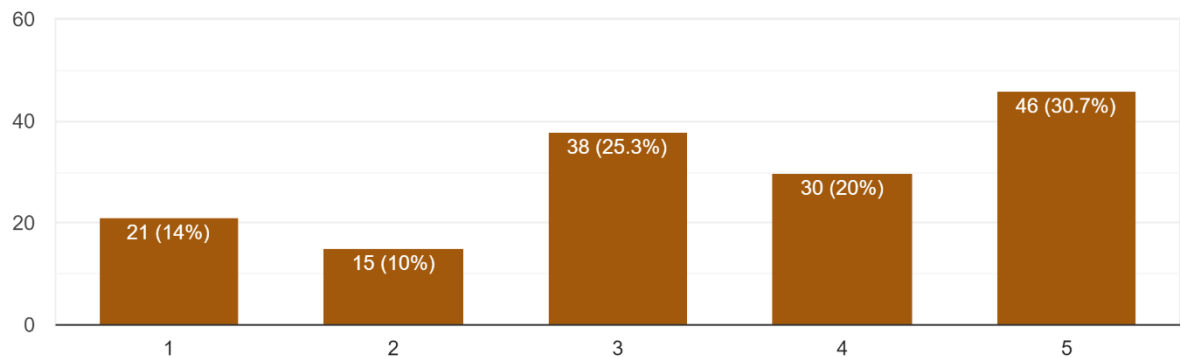


Figure 4.5 - Questionnaire Response 3

Question	On a Scale of 1 to 5 how would you scale your skills in solving Sinhala Arithmetic Problems?
Aim of the Question	To identify if the responder is able to solve arithmetic problems.
Observations	30.7% have rated with the scale of 5, 25.3% have rated with the scale of 3, 20% have rated with the scale of 4, 14% have rated with the scale of 1 and 10% have rated with the scale of 2.
Conclusion	Most of the responders have rated their skills of solving Sinhala Arithmetic Problems high(scale of 4 and 5). Altogether many have rated to be their skills less than 3(scale of 1,2 and 3). Which means Most of the responders have trouble solving Sinhala Arithmetic Problems.

Table 4.5 - Questionnaire Findings 3

How would you normally solve such problems? (eg: වෙළඳසැලක ඇති පරිපූජ්‍ය ප්‍රමාණයෙන් 3/7ක් ඉතිරිවන සේ පළමු මාසය තුළ විකුණන ලදී. පළමු ...දසැල් තිබූ මුළු පරිපූජ්‍ය ප්‍රමාණය කොපමණද?)
150 responses

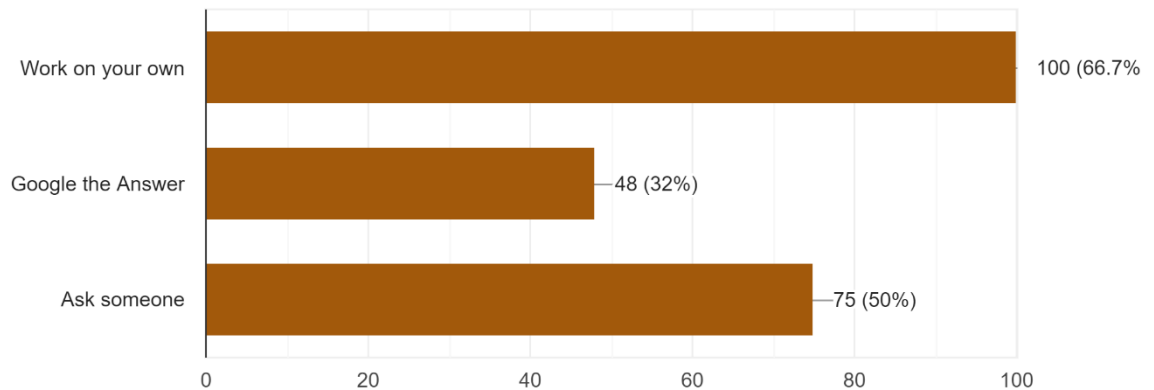


Figure 4.6 - Questionnaire Response 4

Question	How Would You normally solve such problems?(eg : වෙළඳ සැලක ඇති පරිපූජ්‍ය ප්‍රමාණයෙන් 3/7ක් ඉතිරිවන සේ පළමු මාසය තුළ විකුණන ලදී. පළමු මාසය අවසානයේදී විකුණා ඇති පරිපූජ්‍ය ප්‍රමාණය 28kg නම් වෙළඳසැල් තිබූ මුළු පරිපූජ්‍ය ප්‍රමාණය කොපමණද?)
Aim of the Question	To see what the responder will do if they get a Sinhala Arithmetic Problem
Observations	66.7% of the responses have selected 'Work on your own'. 50% have selected 'Ask someone'. 32% have selected 'Google the Answer'.
Conclusion	The responses show that most of them would work on their own and try to solve the problem. While many would ask someone or google the answer. Which means many of them are relying on google and someone else when solving such Sinhala Arithmetic Problems.

Table 4.6 - Questionnaire Findings 4

Do you find these questions to be complicated when you try to solve them?

150 responses

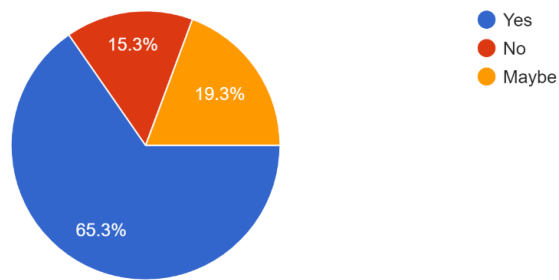


Figure 4.7 - Questionnaire Response 5

Question	Do you find these questions to be complicated when you try to solve them?
Aim of the Question	Get know if the responder's have trouble when they try to solve these types of questions.
Observations	65.3% have answered Yes. 19.3% have answered Maybe and 15.3% have answered No.
Conclusion	Most of the responder's find it difficult to try and solve these kind of Sinhala Arithmetic Problems.

Table 4.7 - Questionnaire Findings 5

Have you ever encountered system that generate answers for Sinhala Arithmetic Problems?

150 responses

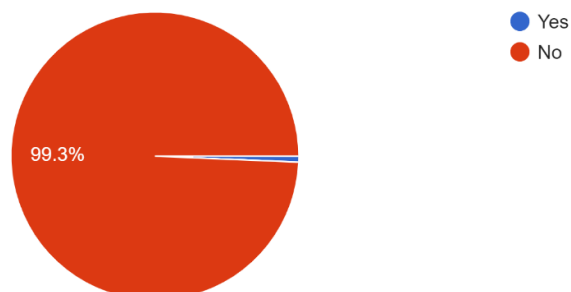


Figure 4.8 - Questionnaire Response 6

Question	Have you ever encountered systems that generate answers for Sinhala Arithmetic Problems?
Aim of the Question	To find if the responder's have come across or have used any sort of systems or applications that generate answers for these types of questions
Observations	99.3% have Answered No and only 0.7%(1 person) has answered Yes
Conclusion	Other than one person all the other responder's have not come across any platform that generate answers for Sinhala Arithmetic Problems

Table 4.8 - Questionnaire Findings 6

If yes, what is the name of that system?

3 responses

N/A

In English, There's an app called Photomath and Google Lens.

Not heard of any

Figure 4.9 - Questionnaire Response 7

Question	If yes, what is the name of that system?
Aim of the Question	To find if there are any related work/ system which the responder might have used.
Observations	Even though only one person has said Yes to the previous question and three have answered this question and only one have given a suitable answer of 'Photomath and Google Lens'

Conclusion	Only one responder have seen a similar system which generates answers for arithmetic problems and it is not available for Sinhala Language.
-------------------	---

Table 4.9 - Questionnaire Findings 7

Do you think such system would be useful to solve Sinhala Arithmetic Problems?

150 responses

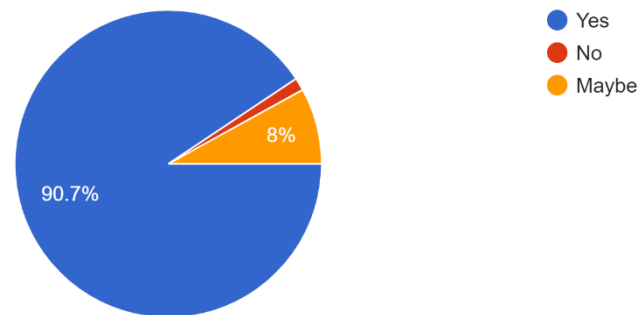


Figure 4.10 - Questionnaire Response 8

Question	Do you think such system would be useful to solve Sinhala Arithmetic Problem?
Aim of the Question	To ask the responder's if a system that would generate answers for Sinhala Arithmetic Problems would be useful.
Observations	90.7% have answered Yes and 8% have answered Maybe and 1.3% have answered No.
Conclusion	Many of the responders would like to see a system that would generate answers for Sinhala Arithmetic Problems.

Table 4.10 - Questionnaire Findings 8

What features would you like to see in this system?

19 responses

- N/A
- User friendliness
- Accuracy
- Voice output of answers and questions
- Ability to convert an problem from english to sinhala
- Maintain a history of problems solved by a user
- Anything
- all mathematical calculations like + - / *
- Child friendly UI

- English translation of the problem
- Steps for the solution
- A system which can recognize human handwriting
- I don't know
- When I add the question to the system it must generate the answer within few seconds
- Speech to text and solve it
- Enter the problem
- Accuracy
- Accurate answers, workings with step by step explanation. And if it could be translated to user preferred language, it would be really helpful for someone like me who can barely read or understand Sinhala

Figure 4.11 - Questionnaire Response 9

Question	What features would you like to see in this system?
Aim of the Question	Get an idea of what kind of features would the responders like to see in the system

Observations	<p>Some responders have not clearly understood the question and have answered, and some answers like ‘English translation’ is out of the scope.</p> <p>There are some same answers as ‘Steps of which the problem is solved’ and ‘speech to text solving’ other than that the answers that are given does not suit for a feature or out of the scope.</p>
Conclusion	Showing the steps of how to solve the problem is a feature that can be looked into and speech to text problem solving could be a future enhancement for the system.

Table 4.11 - Questionnaire Findings 9

4.5.3 Findings from Brainstorming

Criteria	Findings
Discussing the research idea and research gap with supervisor	<p>Different approaches that can be used for the research.</p> <p>Different methods to study and experiment.</p>

Table 4.12 -Findings from Brainstorming

4.5.4 Findings from Observation

Criteria	Findings
How related work have identified different methods to present their work	<p>Identifying research gap and the related work around it.</p> <p>Different requirements the system should have.</p>

Table 4.13 -Findings from Observation

4.6 Summary of Findings

The summary of findings are listed as bellow.

Findings	Literature Review	Questioner	Brainstorming	Observation
Related work and the research gaps available in them.	X			X
Different methods and approaches that has been taken in similar research's	X	X	X	X
The research gap is not have been implemented and only have been identified	X	X		X
Different features that would be useful in the system		X		
Future enhancements that can be done to the system	X	X	X	
Easy to use simple UI implementation		X		

Table 4.14 - Summary of Findings

4.7 Context Diagram

The following describes the context of the system which shows the different components which will be used in the system.

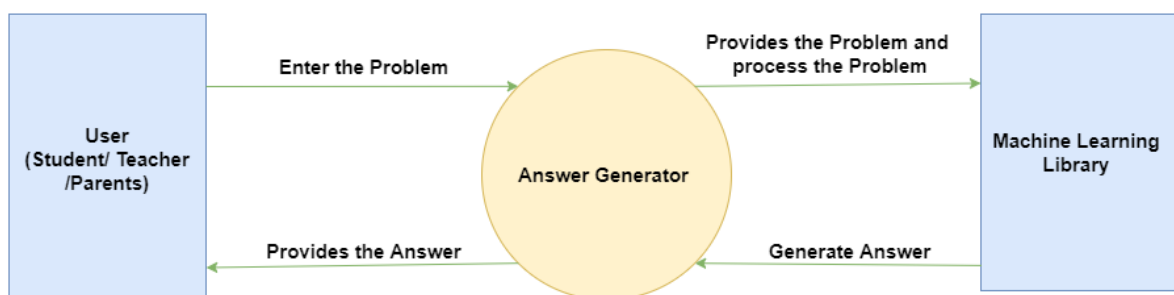


Figure 4.12 - Context Diagram

4.8 Use Case Diagram

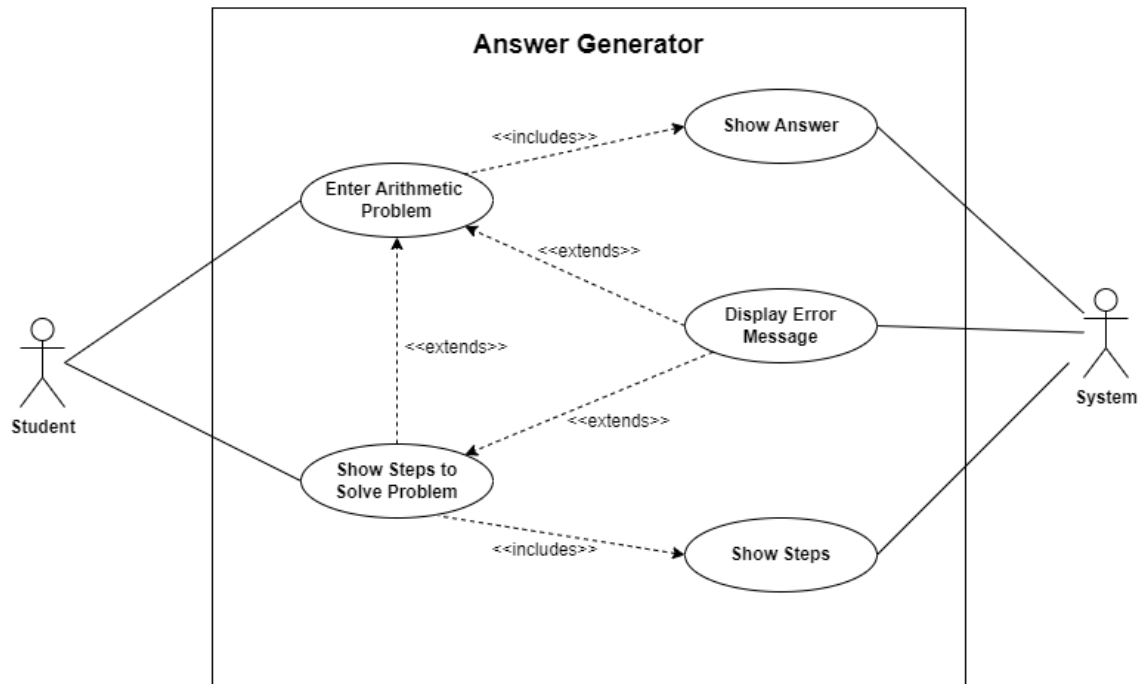


Figure 4.13 - Use Case Diagram

4.9 Use Case Description

Use Case ID	UC-01
Use Case Name	Enter Arithmetic Problem
Description	Student enters the arithmetic problem to the system.
Actors	Student
Pre-Condition	-
Main Scenario	<ol style="list-style-type: none"> 1. Student enters the problem to the system 2. System checks the language if its Sinhala 3. System send the problem to the system. 4. System displays the answer or the error message
Post-Condition	Answer for the entered problem would be displayed

Table 4.15 - Use Case Description UC-01

Use Case ID	UC-02
Use Case Name	Show Steps to Solve Problem
Description	System shows the steps on solving the entered problem.
Actors	Student
Pre-Condition	Student should have entered a problem and the system should have given the answer for it.
Main Scenario	<ol style="list-style-type: none"> 1. Student selects the show steps. 2. System generate the steps on solving the problem. 3. System display the steps on solving the problem to the student.
Post-Condition	Steps on how to solve the problem is shown.

Table 4.16- Use Case Description UC-02

Use Case ID	UC-03
Use Case Name	Show Answer
Description	Displaying the answer of the entered problem
Actors	System
Pre-Condition	Student should have entered a problem to generate the answer.
Main Scenario	<ol style="list-style-type: none"> 1. System gets the entered problem. 2. Identify the keywords in the problem. 3. Execute the model. 4. Generate the answer. 5. Display the answer.
Post-Condition	Answer for the problem would be displayed

Table 4.17- Use Case Description UC-03

Use Case ID	UC-04
Use Case Name	Display Error Message

Description	This will display if there was an error when generating the answer or generating the steps of the problem
Actors	System
Pre-Condition	Student should have entered the problem
Main Scenario	<ol style="list-style-type: none"> 1. System checks for the required output(generate answers or show steps) 2. Identify the error in the system 3. Display the error to the student
Post-Condition	Display the reason for not completing the task given by the student

Table 4.18 - - Use Case Description UC-04

Use Case ID	UC-05
Use Case Name	Show Steps
Description	Displaying the steps on how to solve the problem
Actors	System
Pre-Condition	Student should have entered the problem and should have generated the answer for the problem.
Main Scenario	<ol style="list-style-type: none"> 1. System checks for the entered problem and the generated answer. 2. Runs the model and sperate the keywords in the problem. 3. Generate the steps using keywords. 4. Display the steps.
Post-Condition	Display the steps for the problems.

Table 4.19 - Use Case Description UC-05

4.10 Requirement Specification

Requirements are prioritized according to the bellow table.

Priority Level	Description
Must Have	The systems core functionality
Should Have	It's not very important as the core but will act as a additional function
Could Have	Requirements which can enhance the system but not in the research scope
Will not Have	Requirements that would not be a part of the system

Table 4.20 - Requirement Specification

4.10.1 Functional Requirements

The functional requirements that will be in the system with the priority of each requirement.

Requirement ID	Requirement Description	Priority Level
FR1	The system should generate answers for the arithmetic problems given by the user	Must Have
FR2	The model should be able to identify the keywords and generate the answer	Must Have
FR3	The model should be able to preprocess the data set so that the NLP model to process the data and complete the necessary task	Must Have
FR4	The system should have a classification to classify the arithmetic problems	Should Have
FR5	The system should have a simple easy to use user interface for the user to interact	Should Have

FR6	The system should have a help section so that the user could understand how to use the system	Could Have
FR7	Should be publicly available	Will not Have

Table 4.21 - Functional Requirements

4.10.2 Non-Functional Requirements

The non-functional requirements that will be in the system with the priority of each requirement.

Requirement ID	Requirement Description	Specification	Priority Level
NFR1	The accuracy of the system should at maximum	Accuracy	Should Have
NFR2	The answer generation should be done with less amount of time	Performance	Should Have
NFR3	Adding different features should not affect the core of the system	Maintainability	Could Have
NFR4	The system should not crash while using it	Maintainability	Could Have
NFR5	The system should be able solve complicated arithmetic problems	Scalability	Could Have
NFR6	The system should not be complicated for the user to use	Usability	Could Have

Table 4.22 -Non-Functional Requirements

4.11 Chapter Summary

The chapter started with explaining the system using the rich picture the different stakeholders were identified and reviewed using the onion model. The analysis of requirement elicitation methodologies were identified and discussed. The findings of the questioner and the summary of it was discussed. The context and use case diagrams along with the use case description were given. In the end the functional and non-functional requirements were identified and discussed.

CHAPTER 5 : SOCIAL, LEGAL, ETHICAL AND PROFESSIONAL ISSUES

5.1 Chapter Overview

Discuss in detail using the BCS code of conduct the SLEP issues relevant to your research project and how it was mitigated. You do not have to state the BCS code of conduct explicitly here, however, establish the relationship with the BCS conduct with the SLEP.

5.2 Social, Legal, Ethical and Professional Issues

Social	Legal
<ul style="list-style-type: none"> Distributed questionnaire responses have not been included in the thesis, but only the results were included. Therefore, there's no violation of user privacy. In the survey, none of the user information were collected and it had been utterly anonymous. Users were created privy to the questionnaire being anonymous in its description which by filling the survey, they permit the author to use collected data for analysis purposes. 	<ul style="list-style-type: none"> The data used was retrieved under GPL-3 license and all the tools and languages are under open-source licenses. The dataset that has been used in the project was created by the author. So, there aren't any legal issues with it. GDPR regulations were read and understood and have not been violated by the project or the system produced by the research. Only necessary info was taken, and also the remainder of the unwanted information was safely discarded. Any data stored was stored in a safe manner.
Ethical	Professional
<ul style="list-style-type: none"> Authorships and references were properly cited and respected. 	<ul style="list-style-type: none"> Software Engineering best practices and guidelines of industry were followed

<ul style="list-style-type: none"> • Content which was taken from literature were properly cited and paraphrased in writer's words. Hence no part of the dissertation is plagiarized. 	<p>properly and extensively inside the application.</p> <ul style="list-style-type: none"> • Anonymization of the code used within the project: As good practice, it is imperative that suitable one-way anonymizing functions are applied so that the code can't be re-identified or reproduced. • From the beginning all project code versions have been maintained in the private repository on GitHub with proper commits and tags. • Computers and other data storing systems were password-protected properly
--	--

Table 5.1 - Social, Legal, Ethical and Professional Issues

5.3 Chapter Summary

As mentioned in the requirement elicitation chapter, the BCS Code of Conduct was not violated during the design, development and execution of this research project. All relevant license situations have been adhered to at some stage in the prototype implementation section of this research project.

CHAPTER 6 : SYSTEM ARCHITECTURE & DESIGN

6.1 Chapter Overview

This chapter contains the initial designs of the system. To begin with the design goals are discussed. The high level architecture diagram followed with the component diagram, data flow diagram and system process diagrams are given to describe the designs of the system. Finally, the wireframes UI designs to give the idea of how the user interfaces would look like after the implementation.

6.2 Design Goals

Design goals are discussed in order to get the designs accurately.

Goal	Description
Correctness	The system main objective is to generate answers to arithmetic problems. Natural Language Processing would be the best solution for the system since it would be better than other solutions like deep learning since the correctness is important.
Performance	The system should produce the result in a short amount of time. Since the user is expecting for better performance the best option for it would be to generate the answer for the problem entered.
Scalability	The system's core functionality should work with out any interruptions if the system would get more functions added and scale up the system within the scope.
Reusability	The system should be designed in accordance with having in mind that future enhancements might take place. So, the system should be designed to be reusable for future works.

Table 6.1- Design Goals

6.3 System Architecture Design

Designing high level diagrams gives the basic idea of the system. With the high level designs the layered architecture diagram is given and discussed.

6.3.1 Layered Architecture

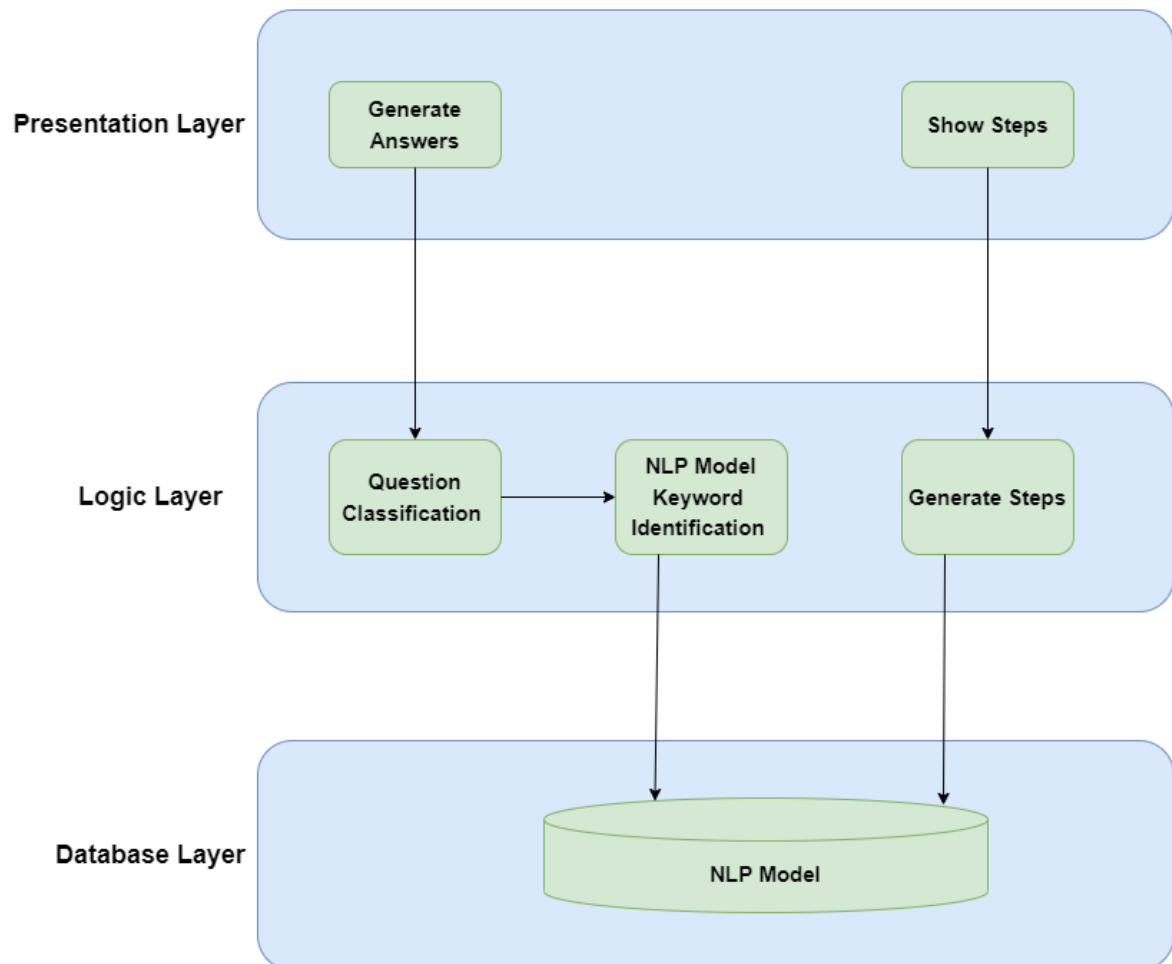


Figure 6.1 - Layered Architecture

The system has three layers which are presentation layer, logic layer and database layer. The presentation layer is the web application of the system or the user interface of the system where the user is interacting, the logic layer is where logic of the system contains with the request which will be given by the user in the presentation layer. The keyword identification and identifying the arithmetic operation is performed in this layer. The database layer is where the NLP model will come into action as the keywords and the identified arithmetic operation will be passed and the generating the answer would be conduct.

Each of the process in the layered architecture are as follows,

- 1) Generate Answers – Getting the arithmetic problem from the user through the user interface(UI).
- 2) Show Steps – Getting the option of show steps from the user.
- 3) NLP Model Keyword Identification - This will check for the main keywords in the problem such as the numerical values.
- 4) Identify the arithmetic operation – This will check for the arithmetic operator(multiply or divide) for the calculation.
- 5) NLP Model – The pre trained model that will help to generate the answer for the problem and the steps to solve it.

6.4 System Design

The system design would give much more idea about the system compared to the layered architecture diagram. The design paradigm would be discussed as well as the different levels of flow diagrams. Also, the activity diagram for the core of the project and the UI diagrams of the system would show on how the user interface would look like.

6.4.1 Design Paradigm

Structured System Analysis and Design Methodology(SSADM) will be used for the system development. Since the research follows the prototyping methodology for development it is suited for design SSADM is used.

Why use Structured System Analysis and Design Methodology (SSADM)?

SSADM helps the research divide into step by step stages where the research begins from feasibility all the way to prototyping. The different stages help the developer identify different requirements to design different diagrams such as high level diagrams, data flow diagrams, system process flowchart and UI designs to understand the designs before prototyping.

The system has a flow where the problem is entered the keywords in the problem should be identified after the arithmetic operator is identified then only the pre trained model would be executed. After one process is completed then will move on to the next process hence SSADM is better suited for the research.

6.4.2 Data Flow Diagram

Data flow diagram shows how the data flow through the system until the system has completed the task given from the user. There are different levels of flow diagram only level-0 and level-1 diagram that to only for the main core of the project is given.

6.4.2.1 Level-1 Data Flow Diagram

The Level-1 dataflow diagram expands the single process of the context diagram into different components which are important to the system. Each of these components are essential for main core of the system that is to generate answer as well as generating step in solving the problem. The diagram shows how each component comes into the action where from the student enters the problem to the answer generation is performed.

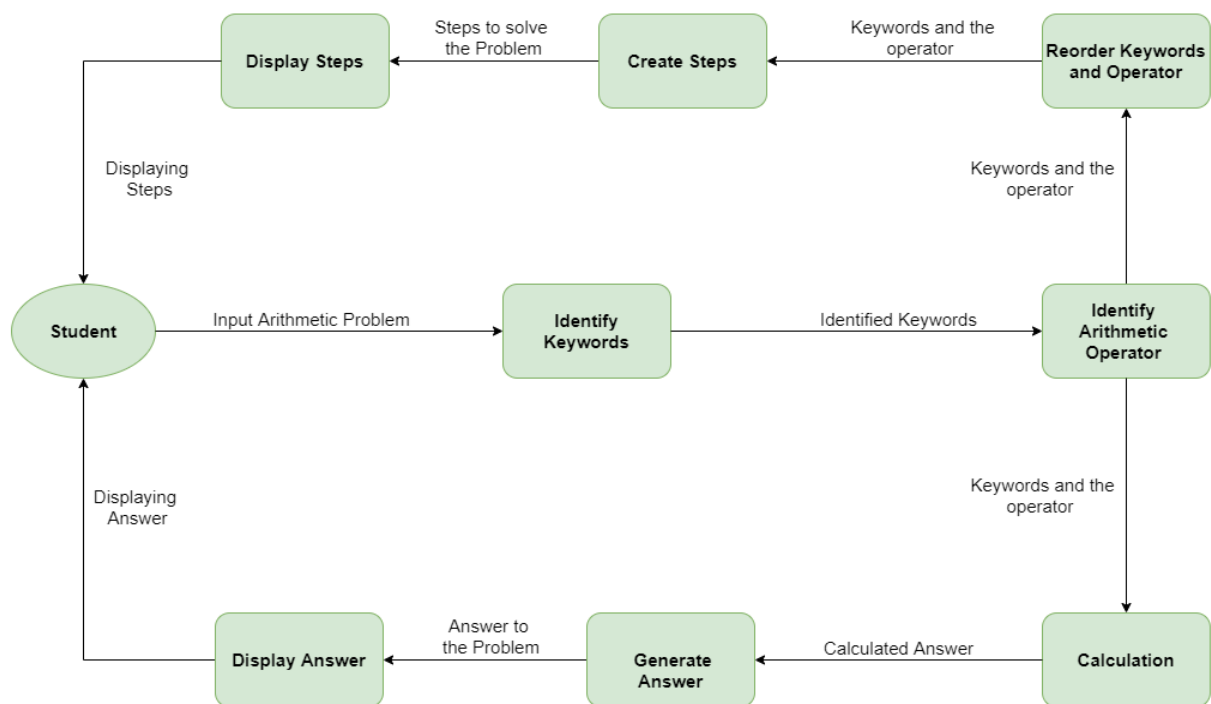


Figure 6.2 - Level-1 Data Flow Diagram

6.4.2.2 Level-2 Data Flow Diagram

The level 2 data flow diagram is created only for the main core of the system in which generating the answer for the entered problem from the user.

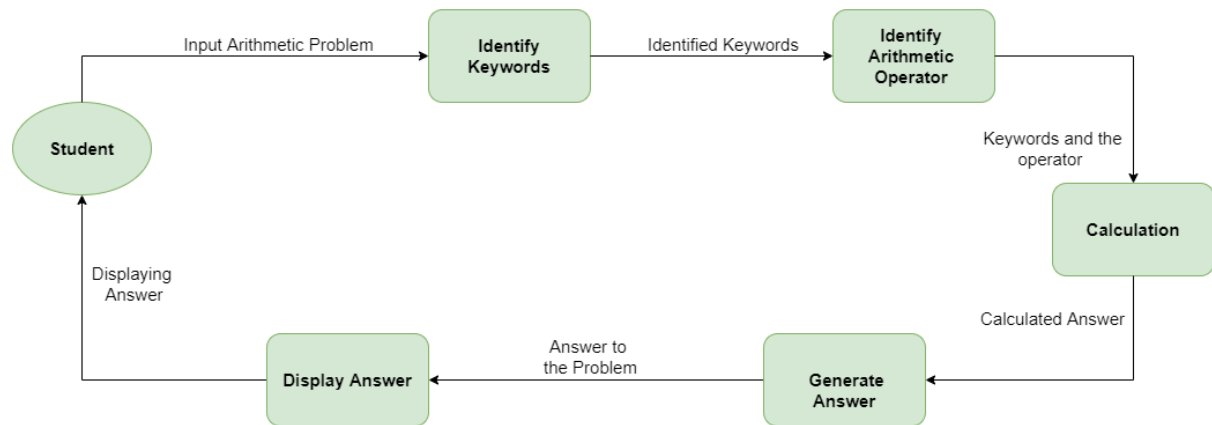


Figure 6.3 - Level-2 Data Flow Diagram

6.4.3 Activity Diagram

The activity diagram is designed only for the main core of the system and it represents how the student and the system is behaving in the process.

Appendix C – Activity Diagram

6.4.4 UI Designs

For UI designs wireframes are designed to give the understanding of the web applications view. The UI designs are important as all the other diagrams represent the system and only the UI diagram represent the how the user will interact with the system.

The diagram starts with the student enters the arithmetic problem through the input field and submit the problem using generate answer, then right next to it the answer for that problem would be displayed. After the answer is generated the student has another option to find out the steps on how to solve the problem with the show steps right under the answer. By selecting the show steps the steps on how to solve the problem would be displayed bellow the problem.

Appendix D – UI Designs

6.5 Chapter Summary

The chapter discussed the different diagrams that would help the implementation. The design goals were identified and discussed. The high level architecture diagram was discussed along with the design paradigms. The data flow diagrams and the system process and UI designs were also discussed that will be used during the implementation.

CHAPTER 7 : IMPLEMENTATION

7.1 Chapter Overview

The chapter would contain the implemented code from the designs that were created, the requirements that were gathered and the literature review that was written all combined from the prototype creation to the final system. The technology selection in this implementation along with the code for the core functionality is discussed. The self-reflection covers the different difficulties encountered during the prototyping and final implementation stage.

7.2 Technology Selections

The technology selection discuss the different technologies used in the implementation such as programming language, libraries, ide's. The technology stack covers the technologies and each one is discussed below.

7.2.1 Technology Stack

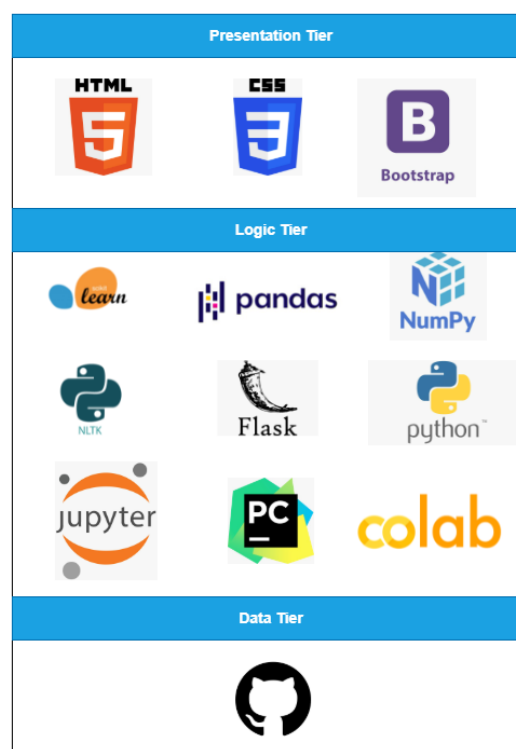


Figure 7.1 - Technology Stack

7.2.2 Data Selection

The research is a data science research due to that a having a dataset is a must. Since the domain is a less researched area since the Sinhala language is less resource domain. The author had trouble finding a correct dataset which suits the research. Kaggle was the first place the author looked for the dataset there was nothing similar which could help in the research. Open data portal Sri Lanka was the next best platform to find datasets in which was unsuccessful, since the needed dataset was not found there also. So, the author had no choice but to create a dataset which suits the research. The data was taken from grade 5 Sinhala medium mathematical textbook. The data set was created and had over 200 records at the time.

7.2.3 Programming Language

Considering Java and Python out of those programming languages python was selected to implement the model. The reason for it would be python support many libraries and easy to work with and most of the related work has shown their approaches have been achieved using python. Also, python has a very well document when using its libraries such as pandas, scikit learn and many.

- Python is interpreted programming language where code is executed line by line.
- Python is free and open source where its easy to use and it does not cost anything.
- Many libraries are available in python that helps built the project

7.2.4 Libraries

All the libraries which has been used to implement the system are discussed here.

Library	Description	Features
Pandas	Pandas was used to manipulate the dataset it helps to get analsis of the data set and manipulate it however the model requires.	<ul style="list-style-type: none"> • Clean up datasets • Visually display dataset • Support different formats of datasets
Scikit Learn	Scikit learn is very popular library in python that can be helpful in classification, regression, and clustering python models.	<ul style="list-style-type: none"> • Split data for training and testing

		<ul style="list-style-type: none"> • Easy to create linear regression
NumPy	NumPy is popular library among python developers that helps when working with arrays and metrics.	<ul style="list-style-type: none"> • High performance • Multidimensional container
NLTK	NLTK is python library which supports NLP. This helps preprocess data so that it is easy for the dataset to be used in the model for training and testing.	<ul style="list-style-type: none"> • Tokenizing • Part of speech tagging • Sentiment analysis

Table 7.1 - Libraries

7.2.5 Selection of Development Framework

Since the model of the project is developed using python the requests that would come from the frontend and the final output of the model is delivered to the frontend using flask framework.

Why use Flask?

Flask is easy to use not so complicated backend framework which can connect frontend framework and the python backend with ease. The author have worked with flask and found out its easy to use and support any frontend framework when it comes to sending or receiving requests from the backend. Rather than having a much complex backend flask is a simple and more efficient comparing to other frameworks.

Framework	Benefits	Drawbacks
Flask	<ul style="list-style-type: none"> • Support every version of python without any errors • Well documented and community support is high • Easy to use with frontend frameworks as well as simple frontends. 	<ul style="list-style-type: none"> • Takes time to execute compared to other frameworks

Table 7.2 - Why use Flask

7.2.6 Selection of IDE's

In order to develop the coding of the system an ide should be used. Comparing with many ide's which are available the author has used Jupyter Notebook and PyCharm. Jupyter notebook was the starting editor when initially the implementation started as it's easy to use. When it came to developing the whole system, the notebook is not capable of archiving it so PyCharm was used during the next stages of implementation. PyCharm supports all the libraries and also it helps to identify any errors in the code before execution.

7.2.7 Selection of Version Control

It is essential that when developing a project, the developer should keep a version of codes as sometimes the code might get bugged and cannot resolve and need to go back and get a older version of it. For this project the author has used GitHub as the version control platform as it is easy to commit to git repository using PyCharm itself. And GitHub will store all the version of commits without a single data loss.

Version Controller	Benefits	Drawbacks
GitHub	<ul style="list-style-type: none"> Keep a track of the version control in the code GitHub can be intergraded with PyCharm as it is easy to commit changes Let's you store your work without any drawbacks 	<ul style="list-style-type: none"> Repositories can get security issues that would cause loose some of the data in the repository

Table 7.3 - Selection of Version Control

7.2.8 Summary of Technology Selection

The summary of technologies used in each of the different areas are listed below.

Programming Language	Python
Libraries	Pandas, Scikit Learn, NumPy, NLTK
Backend Framework	Flask

IDE	Jupyter Notebook, PyCharm
Version Control	Git Hub

Table 7.4 - Summary of Technology Selection

7.3 Core Functionality

The section would discuss the answer generation for the arithmetic problem how it is achieved from start to finish.

7.3.1 Data Preprocessing

For the project two datasets were created one for the training and testing purpose for the classification the other dataset to test the entire system. The classification dataset contains the question and the category of the question it falls into.

	Question	Label
0	කණුවක බැඳුන් බෝල 2 ක බැගින් බැඳුන් බෝල 600 ක් ...	Division
1	බිස්කට් 300 ක් ඇත. ළමයෙකුට බිස්කට් 2 බැගින් ළම...	Division
2	මීටර 90 ක් දිග කම්බියක් ඇත. සමාන කැබලි 2 බැගින්...	Division
3	මීටර 60 ක් දිග ඉටි රෙද්දක් ඇත. මෙසවලට දැමීම සද...	Division
4	පැන්සල් 420 ක් ඇත. ළමයෙකුට පැන්සල් 2 බැගින් ළම...	Division
...
324	කේට් ඇපල් 154 ක් මිලදී ගත් අතර එයින් 83 ක් කුණ...	Subtraction
325	වේලර් විසින් අතින් සාදන ලද භාණ්ඩ 66 ක් දේශීය ප...	Subtraction
326	පැසිපන්දු තරගයකදී ගෝඨා පිල ලකුණු 50 ක්ද රනිල් ...	Subtraction
327	රොන් පරීක්ෂණයෙන් ලකුණු 76 ක් ලබා ගන්නා අතර ඔහු...	Subtraction
328	ළමයින් 45 දෙනෙකුගෙන් යුත් කණ්ඩායමක් රචනා තරඟයක...	Subtraction

Figure 7.2 - Label Dataset

The dataset which is used for the project is a 2 columns based on the question column and the label for the question. The questions have their own answers which are different to each of the data record. The dataset is still under creating to make sure it has enough data when it comes to training and testing.

	Question	Answer
0	බිස් රථයක මගීන් 58 දෙනෙකුට ගමන් කළ හැකි ය. එවැනි...	14384
1	පැනක මිල රුපියල් 12 කි. එවැනි පැන් 19 ක මිල කො...	228
2	මාර් බිස්කට් පැකට් එකක මිල රුපියල් 100 කි. පෙට...	2200
3	කෙසෙල් ගෙඩි 840 ක් ඇත. කෙසෙල් ගෙඩි 35 ක් බැගින්...	24
4	පැනක මිල රුපියල් 12 කි. එවැනි පැන් 37 ක මිල කො...	444
...
455	බිස්කට් 300 ක් ඇත. ළමයෙකුට බිස්කට් 100 බැගින් ...	3
456	කුඩයක නාරං ගෙඩි 2 ක් ඇත. එවැනි කුඩ 22 ඇති මුළු...	44
457	කණුවකට බැඳුන් බෝල 50 ක් ඇත. බැගින් බැඳුන් බෝල ...	12
458	අඹ ගෙඩි 1200 ක් ඇත. අඹ ගෙඩි 8 බැගින් පෙට්ටි වල...	150
459	මකනයක මිල රුපියල් 20 කි. පොතක මිල මකනයක මිල ම...	240

Figure 7.3 : Normal Dataset

During the preprocessing stage all the unwanted characters such as html markups, punctuations.

```
def preprocess_and_tokenize(data):
    #Remove html markup
    data = re.sub("<.*?>", "", data)

    #Remove urls
    data = re.sub(r'http\S+', '', data)

    #Remove hashtags and @names
    data= re.sub(r"#[\d\w\.]+)", '', data)
    data= re.sub(r"@[\d\w\.]+)", '', data)

    #Remove punctuation and non-ascii digits
    data = re.sub("(\\W|\\d)", " ", data)

    #Remove whitespace
    data = data.strip()

    #Tokenization with nltk
    data = word_tokenize(data)

    #Stemming with nltk
    porter = PorterStemmer()
    stem_data = [porter.stem(word) for word in data]

    return stem_data
```

Figure 7.4 - Data preprocessing

Then after the dataset was divided into two files one for training and testing. The dataset was loaded to the project using pandas. After the dataset was load the training and testing data splitting were done then the two categories which are “Multiplication” and “Division” were declared in the class array then the data frame was created.

```

model_train = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/Label_Questions_Train.csv')
model_test = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/Label_Questions_Test.csv')

X_train = model_train.Question
X_test = model_test.Question

y_train = model_train.Label
y_test = model_test.Label

class_names = ['Multiplication', 'Division', 'Addition', 'Subtraction']
data = pd.concat([model_train, model_test])

```

Figure 7.5 - Assigning training and testing data

7.3.2 Data Vectorization

For the vectorization TFIDF vectorizer was used. This would calculate the worth of a phrase in a sentence. Merely multiplying two factors gets the result. the chance that a user will enter a sentence and the opposite document frequency over a set of documents.

```

# TFIDF, unigrams and bigrams
vect = TfidfVectorizer(tokenizer=preprocess_and_tokenize, sublinear_tf=True, norm='l2', ngram_range=(1, 2))

# fit on our complete corpus
vect.fit_transform(data.Question)

# transform testing and training datasets to vectors
X_train_vect = vect.transform(X_train)
X_test_vect = vect.transform(X_test)

```

Figure 7.6 - Data vectorization

7.3.3 Classification

In order to perform the question classification Logistic Regression was used. Lbfgs (Limited memory Broyden Fletcher Goldfarb Shanno) solver was used to train the model. A pipeline was created for the vectorizer and logistic regression model. The pipeline was used to stop the accuracy of the model been leak. After the model was saved to pickle file to develop them system in the later part.

```

#Classification using LogisticRegression
log = LogisticRegression(solver='lbfgs', multi_class='auto', max_iter=200)
log.fit(X_train_vect, y_train)

ylog_pred = log.predict(X_test_vect)

```

Figure 7.7 - Training the regression model

```
#Create pipeline with our tf-idf vectorizer and Logistic Regression model
#To stop the leaking of the accuracy
log_model = Pipeline([
    ('lbfgs', vect), #parameter value for the vectoriser
    ('clf', log),    #parameter for the classification
])
```

Figure 7.8 - Creating the pipeline

7.3.4 Answer Generation

The next phase is to generate the answer to the question. The question is tokenized using the sinling Sinhala tokenizer. After the POS tagging was done to identify the tags of the question and to identify the keywords. In this case the keyword the system is looking for is the numerical values in the question.

```
# Tokenization with sinling SinhalaTokenizer ('https://github.com/yusenarath/sinling')
tokenizer = SinhalaTokenizer()
tokenized_sentences = [tokenizer.tokenize(f'{ss}.') for ss in tokenizer.split_sentences(question)]

tagger = POSTagger()
pos_tags = tagger.predict(tokenized_sentences)

# Identifying the numbers in the question and assigning them into the variable
keyword_numbers = [(word, tag) for word, tag in pos_tags[0] if (tag == 'NUM')]
```

Figure 7.9 - Tokenization and POS tagging

Using chunking the necessary keywords are identified. In the POS tagged array the “NUM” tagged keywords are searched and the keywords are being assigned to a variable.

```
# Identifying the numbers in the question and assigning them into the variable
keyword_numbers = [(word, tag) for word, tag in pos_tags[0] if (tag == 'NUM')]
```

Figure 7.10 - Keyword Extraction

As the question that is been passed is a string the extracted keywords are also in have the data type sting in order perform arithmetic calculation the keywords have been converted to integer. Then according to the prediction from the loaded model the category of the question the mathematical calculation is performed.

```
filename = 'Model1.sav' # Model file

model = pickle.load(open(filename, 'rb')) # loading the model
category = model.predict([question])[0] # Predicting the category of the question
```

Figure 7.11 - Predicting the question category

```

num_1 = keyword_numbers[0][0]
num_2 = keyword_numbers[1][0]

number_1 = int(num_1)
number_2 = int(num_2)

print("num 1 : ", num_1, " num 2 : ", num_2)

if category == 'Multiplication':
    answer = number_1 * number_2 # calculating the answer

elif category == 'Division':
    if number_1 > number_2:
        answer = number_1 / number_2 # calculating the answer

    elif number_2 > number_1:
        answer = number_2 / number_1 # calculating the answer

```

Figure 7.12 - Answer Calculation

After the calculation is performed the answer for the question is generated.

7.4 Implementation of API's

Flask was used to create the API's for the system. A flask project was created where the API's along with the user interfaces were also in the project structure.

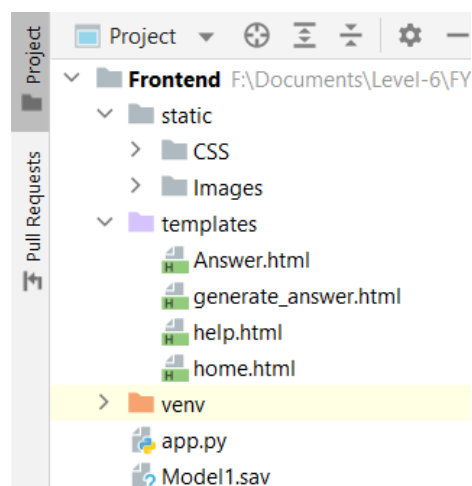


Figure 7.13 - Project structure

There are four templates for the system all the template rendering has been executed in the “app.py” file. The file contains different routes to different pages and the core functionality route is given below to understand the structure.

```
app = Flask(__name__)
app.secret_key = "question_answer" # secret key for session

# Route for the home page
@app.route('/')
def home(): # put application's code here
    return render_template("home.html")

# Route for the generate answer page
@app.route('/generate_answer', methods=["POST", "GET"])
def generate_answer():...

# The preprocessing method of the model
def preprocess_and_tokenize(data):...

# Route for the answers page
@app.route('/answer', methods=["POST", "GET"])
def getanswer():...

# Route for the help page
@app.route('/mathq_help')
def mathq_help():
    return render_template("help.html")

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

Figure 7.14 - API routes

7.5 User Interfaces

The UI is developed in the flask project as templates using HTML and CSS. Below is the landing page or the home page of the system.

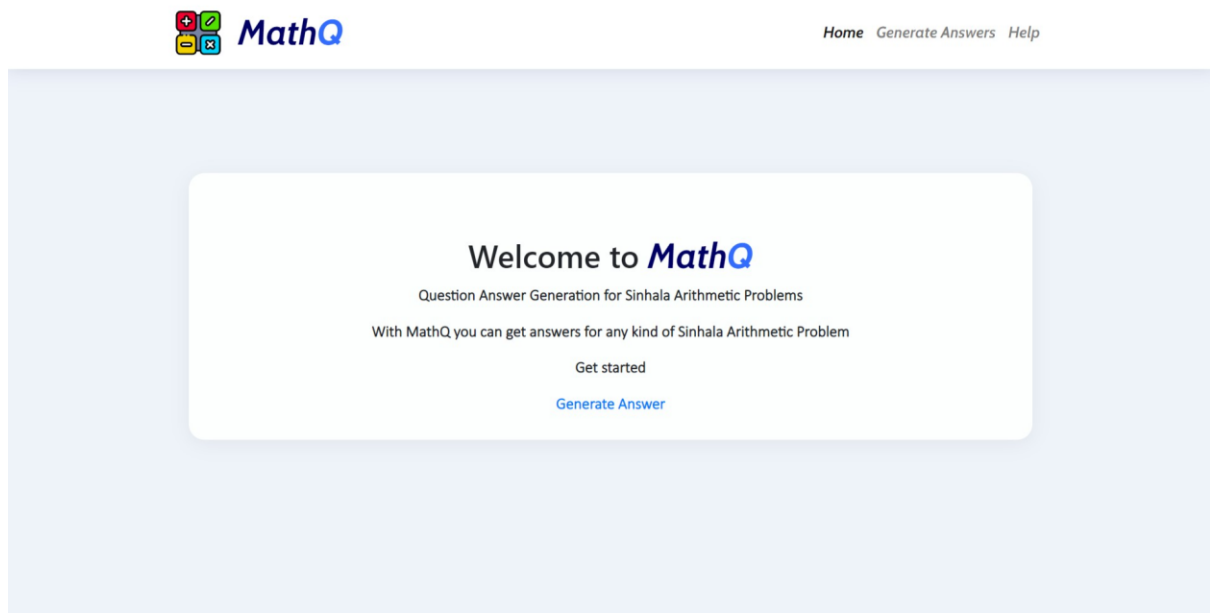


Figure 7.15 - Home page

To access the core of the project user, have to navigate to the Generate Answer page.

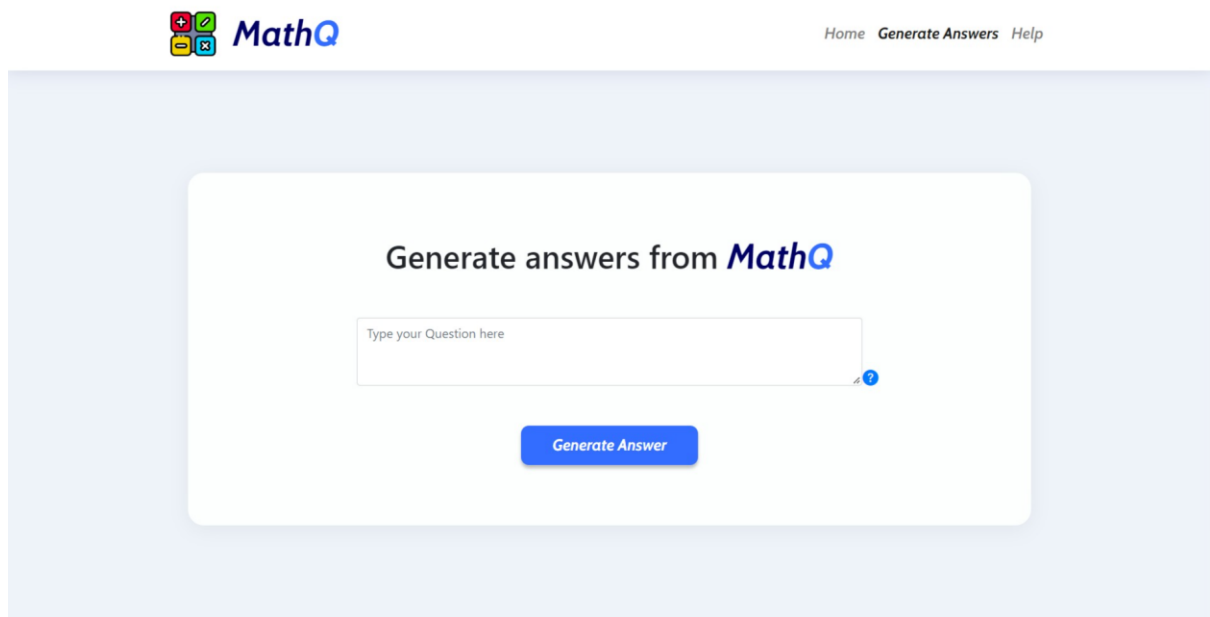


Figure 7.16 - Generate Answer Page

The question should be entered to the text area given and click the Generate Answer button to generate the answer for the question.

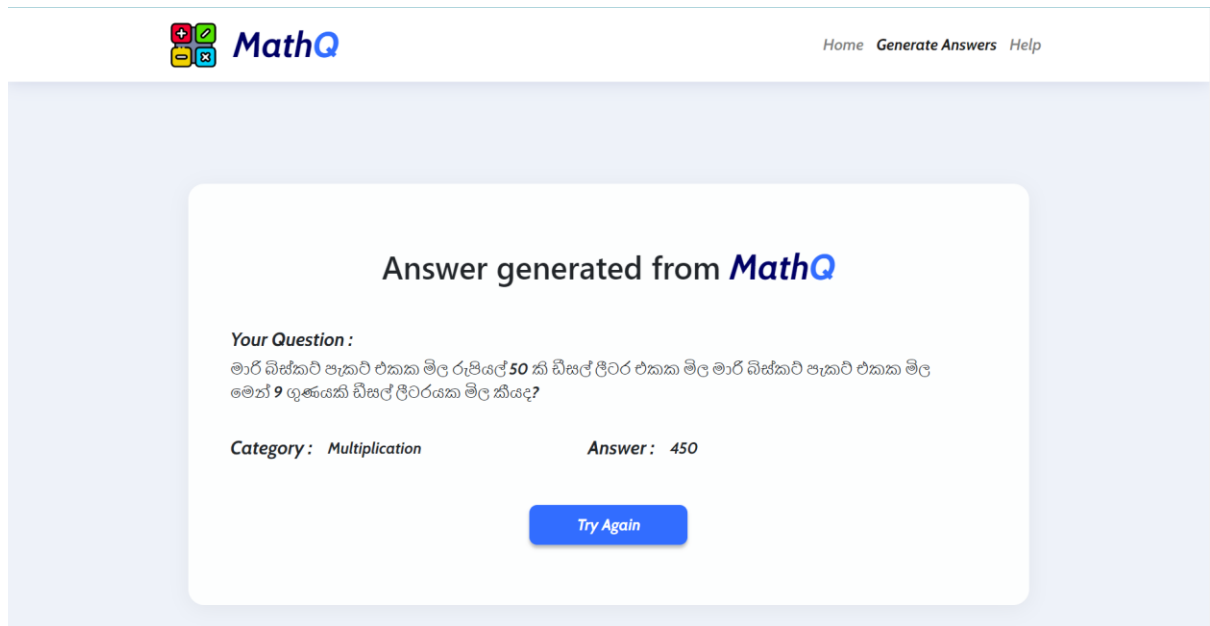


Figure 7.17 - Answers Page

As shown in the above figure the answer for the question the user entered and the category of the question is generated by the system and displayed here.

7.6 Chapter Summary

The chapter covered the different technologies that was used during the implementation and their uses along with the technology stack. The core functionality of the system is discussed and how it is been implemented to get the final output and finally the self-reflection covered the difficulties encountered during this stage.

CHAPTER 8 : TESTING

8.1 Chapter Overview

This chapter contains the testing that was conducted in the implemented prototype. The objectives and goals of testing are discussed and the testing criteria is given. The model testing was conducted and the results are discussed. The functional and non-functional requirements testing is described.

8.2 Objectives and Goals of Testing

The main objective of the testing the implementation is to check whether the functional requirements which was identified previously in the research. Also, the following set of goals were set to check if the requirements of the project has fulfilled.

- The functional requirements are met and implemented in the final prototype.
- The non-functional requirements are met and implemented in the final prototype.
- Check for bugs and errors in the final implementation.
- To identify what needs to be done to get a high accuracy.

8.3 Testing Criteria

The following criteria are developed after first considering the requirement for a criterion to narrow the gap between the proposed solution and the final implementation.

- Functional testing: the core functionality of the system is tested, with the functional requirements that was identified during the start of the project.
- Non-functional testing – the accuracy of the system, the performance of the system are tested.

8.4 Model Testing

Model testing section covers the testing which was conducted on the model. For testing confusion matrix, accuracy of the model, F1 score, precision and the recall have been generated.

8.4.1 Confusion Matrix

The confusion matrix compares the label values in the dataset with the predicted value from the classification model.

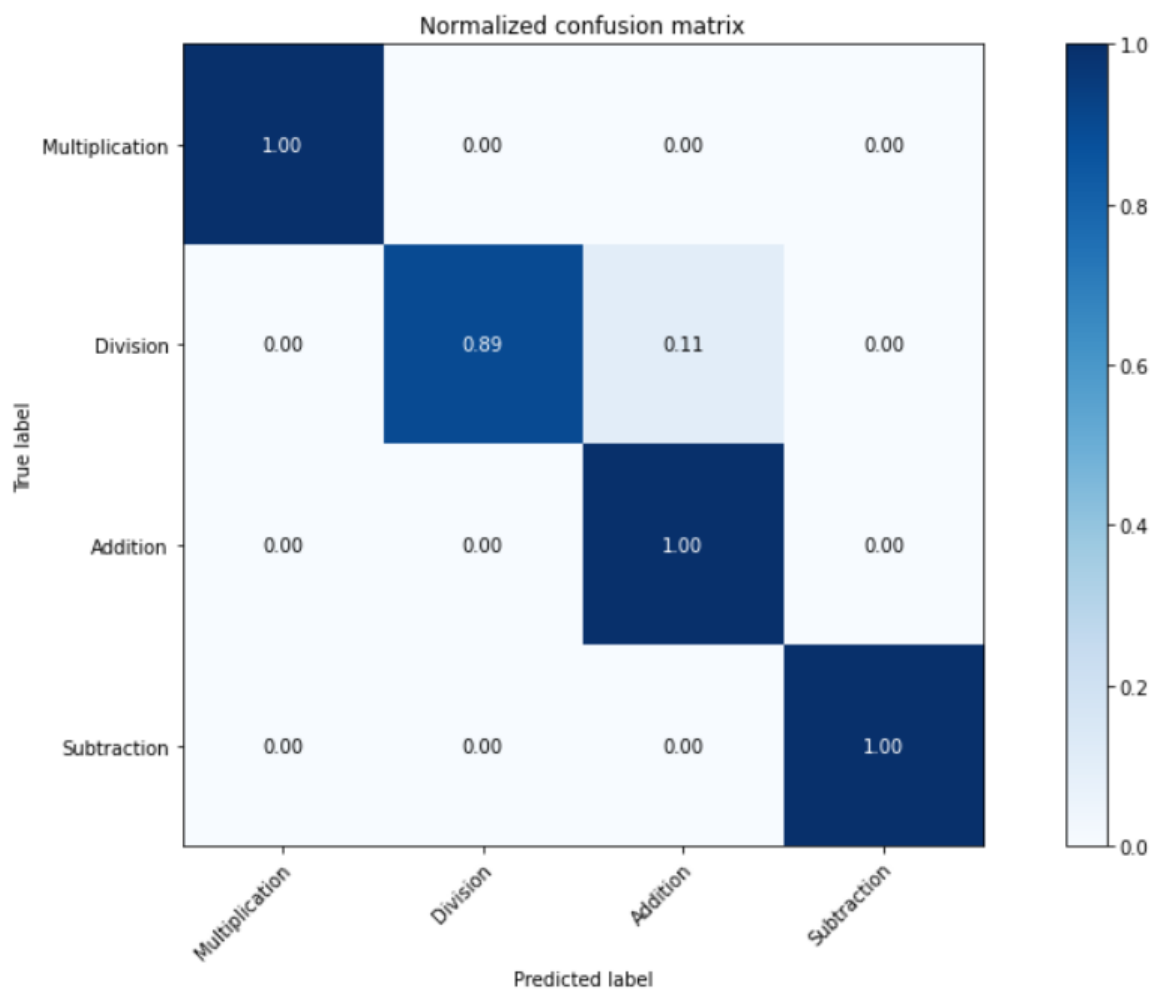


Figure 8.1 - Confusion Matrix

The plotted results show how the model has predicted the label or the category of the question with dataset's actual label value.

8.4.2 Accuracy, F1 Score, Precision, Recall

The accuracy of the model, the F1 score, precision score and the recall score of the model has been obtained using their respective coding and the results have been printed.

```
Accuracy: 97.65%
F1 Score: 97.65%
Recall Score: 97.65%
Precision Score: 97.65%
Confusion Matrix :
[[24  0  0  0]
 [ 0 17  2  0]
 [ 0  0 20  0]
 [ 0  0  0 22]]
```

Figure 8.2 - Accuracy, F1 Score, Precision, Recall

As per the results the classification model has an accuracy of 94.87%. According to that the classification model is able to predict the category of the question with an accuracy of 94.87%. as well as for each of the score F1, Recall and precision the same value of 94.87% has been generated.

```
Number of Correct Answers : 436
Total Number of Questions : 460
System Accuracy : 94.78260869565217
```

Figure 8.3 - System Accuracy

The above system accuracy was calculated with using the dataset which contains both question and answer. In this case the system accuracy is the combination of both classification model as well as the answer calculation of the problem.

8.5 Benchmarking

In this section benchmarking would be conducted with existing work. Comparing MathQ work with the existing work. Only one related work was there in the same domain the Sinhala domain and only the accuracy of the model was mentioned in the research paper.

Research	Accuracy of the model	F1-Score
MathQ (Author's research project)	97.65%	94.87%
Mahoshadha 2(Chathurika et al., 2016)	72%	-
MathQA (Schubotz and Scharpf, 2018)	80%	-

Table 8.1 - Benchmarking

8.6 Functional Testing

For functional testing a black box testing was conducted which the expected outcome and the actual outcome is done for each of the functional requirements.

ID	Input	Expected Outcome	Actual Outcome	Status
FR1 : The system should generate answers for the arithmetic problems given by the user				
FT1	Sinhala arithmetic question	Correct answer for the question	Correct answer for the question	Pass
FT2	English arithmetic question	Display error message	Display error message	Pass
FR2 : The model should be able to identify the keywords and generate the answer				
FT1	Sinhala arithmetic question	Identify keyword in the question	Keywords identified	Pass
FR3 : The model should be able to preprocess the data set so that the NLP model to process the data and complete the necessary task				
FT1	Sinhala arithmetic question	Remove unwanted characters from the sentence	Unwanted characters were removed	Pass
FR4 : The system should have a classification to classify the arithmetic problems				

FT1	Sinhala arithmetic question	Classify the category of the model correctly	The category of the question generated correctly	Pass
FR5 : The system should have a simple easy to use user interface for the user to interact				
FT1	Sinhala arithmetic question	The question is passed to the backend without errors	The question is passed to the backend without errors	Pass

Table 8.2 - Functional Testing

8.7 Module and Integration Testing

Module	Input	Expected Outcome	Actual Outcome	Status
Input text area	Sinhala arithmetic question	Question gets submitted	Question gets submitted	Pass
	Non-Sinhala arithmetic question	Question gets submitted but an error message gets displayed	Error message gets displayed	Pass
Model classification	Sinhala arithmetic question	Predict the correct category of the question	The correct category of the question is predicted	Pass
Keyword extraction	Sinhala arithmetic question	Identify the correct keywords and get extracted	The correct keywords get extracted	Pass

Answer calculation	Sinhala arithmetic question	The correct answer gets calculated	The correct answer is calculated	Pass
--------------------	-----------------------------	------------------------------------	----------------------------------	------

Table 8.3 - Module and Integration Testing

8.8 Non-Functional Testing

This section focuses on the tests run for the non-functional requirements defined in the requirement specification chapter. These non-functional studies are carried out in accordance with the following requirements.

8.8.1 Performance Testing

The system UI is a web application hence it is important to check for the performance of the website. In order to increase the performance of the website the author has test and train the model in Google colab. After successfully training the model the model was saved into a pickle file and later loaded into the system. This would help the classification process work fast rather than executing the whole code. Also, the rest of the code is simple and don't take much time to execute. This helps to increase the performance of the website.

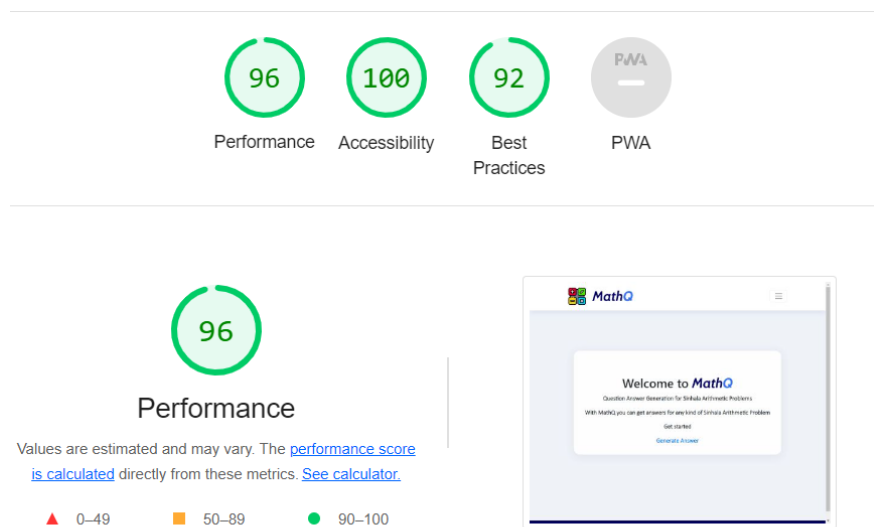


Figure 8.4 - Performance testing

8.8.2 Scalability

For the project the author had to create the dataset from scratch hence the dataset is not enough at the moment. For scalability if the dataset is increased the accuracy of the model would definitely increase as it would help the model to identify new questions which were not their initial implementation. Increasing the dataset would not reduce any performance on the website but increase the systems accuracy.

8.9 Limitations of the Testing Process

During the testing phase the author encountered some limitation. Some of the limitation that was encountered are listed below.

- Not enough data in the dataset due to creating data took a lot of time the author had limit the amount of data in the dataset.
- Since the project is hosted locally the performance testing was limited to it.

8.10 Chapter Summary

This chapter covered the testing for the model in the prototype along with the objectives and goals of testing were identified. The testing criteria's were discussed with the testing results that was conducted. Functional testing and non-functional testing were conducted on the implemented system.

CHAPTER 9 : EVALUATION

9.1 Chapter Overview

This chapter is based on feedback obtained by the associated stakeholders and domain experts, however the many aspects together with determinations made within the system were evaluated in this chapter. This chapter also contains the evaluation criteria, the authors self-evaluation, Limitation to evaluation, Functional requirement evaluation and Non-functional requirement evaluation.

9.2 Evaluation Methodology and Approach

As the main evaluation methodology, the qualitative approach was used. Due to the pandemic situation, the explanation and the demonstration of the system and how it works were explained to the domain experts via online meetings and phone calls. These online meetings were conducted separately for each domain expert in order to gain in-depth knowledge of the system. In order to gather feedback from the domain experts, google forms and emails were used. To make it easier, all of the evaluations have been grouped into a set of criteria.

9.3 Evaluation Criteria

The following criteria were considered in the evaluation method to include the significant stages of the project.

Criteria	Purpose to evaluate
Project concept	To seek feedback and observation on the project from the targeted audience.
Scope of the Project	Defined for the researched domain and problem where a reasonable amount of research considerable depth was included in the research.

Implementation of the Project	To ensure if the project achievements came up to an expected standard of research.
Context of the Thesis	To ensure the research documentation is up to standard and the context of it is correct.

Table 9.1 - Evaluation Criteria

9.4 Self-Evaluation

Author has performed his own Evaluation. Self-Evaluation has done based on to define the evaluation criterions in the section.

Criteria	Self-Evaluation
Project concept	This gives the perfect idea about the project idea and its depth
Scope of the Project	The scope tells the boundaries of the project and most probably lets you identify future enhancements as well.
Implementation of the Project	The implementation has been met the criteria's of the project and the requirements.
Context of the Thesis	The context of the thesis very helpful and very informative.

Table 9.2 - Self-Evaluation

9.5 Selection of the Evaluators

The evaluators were selected under three categories which are,

- Domain experts who have both knowledge and experience in teaching.
- Technical experts who have the experience in the NLP/Machine Learning area.
- Beginners who have less experience in the domain and the technical side.

All of the evaluation feedback were taken from emails after successful meetings with each of the evaluators.

Group	Name	Designation
Domain Experts	Mrs. Sumithra Pasqual	Sri Lanka teaching service grade 1 - Mathematics O/L Kegalu Vidyalaya
	Mr. Nishantha Perera	Mathematics teacher - Gokaralla Vidyodana Piriwena
Technical Experts	Mr. Nisal Suranaka	BEng(Hons) Computer Science - University of Westminster, UK. Software Engineer at Sysco LABS
	Ms. Akshina Fernando	Reading Master of Computer Science, Flinders University, SA. BEng(Hons) Software Engineering - University of Westminster, UK. Former Software Engineer, DMS Software Engineering Pvt Ltd
Beginners	Mr. Arosha Mendis	Mechanical Engineering Undergraduate at SLIIT
	Mr. Gayan Jayawardana	Software Engineer Intern at Omobio

Table 9.3 - Selection of the Evaluators

9.6 Evaluation Result

After presenting the project idea and the implementation of the project to each expert and their feedback of it is listed in this section.

9.6.1 Expert Opinion

9.6.1.1 Project Context

Question	Please provide your thoughts on answer generation for Sinhala arithmetic problems.
Evaluator	Feedback
Mrs. Sumithra Pasqual	The research is a good research area as this would help a lot of students as in the future this would be helping as the students gets involved with technology and finding answers for problems they encounter while studying.
Mr. Nishantha Perera	The idea is good but not practical as students should have to develop their problem solving skills, but this will help the students who are struggling to solve problem on their own to have a helping start with this system.
Mr. Nisal Suranaka	The whole idea of generating answers for Sinhala arithmetic problem looks easy but as a technical expert I see the depth of this project and the complexity of the project hence the project is a good way to test your skill and develop your knowledge in machine leaning/ NLP.
Akshina Fernando	The project idea is a good research as the domain is Sinhala. Sinhala domain is still upcoming domain in researchers in the industry as well. Question answering has a wider research as there are many research based on that, creating a question answering system for Sinhala domain is a better project in my opinion.
Mr. Arosha Mendis	The project is very helpful as the students would have great opportunity to solve difficult mathematical problems in Sinhala language.
Mr. Gayan Jayawardana	The research is much helpful to the society as this would help students solve mathematical problems easily without any hesitation.

Table 9.4 - Project Context

9.6.1.2 Project Scope & Depth

Question	Do you believe this research project has the depth required for a bachelor's degree?
Evaluator	Feedback
Mrs. Sumithra Pasqual	Yes, The research has good depth that will cover for your bachelor's degree.
Mr. Nishantha Perera	Indeed, this research has a wide scope and this has the requirements you need for your bachelor's degree.
Mr. Nisal Suranaka	Absolutely, the project contains a good depth which requires for a bachelor's degree in my opinion. As the title might be seen as a project with simple task but in the industry level this has a good value.
Akshina Fernando	Answer generation for Sinhala arithmetic problems have a good depth as Sinhala domain is less researched area and it is still developing area. So, this research in my opinion has depth which is required for a bachelor's degree.
Mr. Arosha Mendis	Yes, I believe this project has the requirements for it.
Mr. Gayan Jayawardana	Yes, the project scope has wide coverage that would cover for a bachelor's degree.

Table 9.5 - Project Scope & Depth

9.6.1.3 Implementation of the Prototype

Question	Can you give your opinion on implemented prototype?
Evaluator	Feedback
Mrs. Sumithra Pasqual	The prototype is well implemented the answers are also correct. The system is very accurate on generating answers for mathematical problems.
Mr. Nishantha Perera	Each of the questions that was added gave out correct answers. Overall, the system is excellent have not seen such systems before.

Mr. Nisal Suranaka	A simple UI where any user could use, the implementation has met the project objective as the final output is correct. If the dataset has more records of data then the accuracy can be higher.
Akshina Fernando	To train and test the model could have use more complex arithmetic problems but as per the scope the implementation is better. The frontend is very user friendly. The basic arithmetic problems are being identified correctly and the answer that is been generated is correct.
Mr. Arosha Mendis	The system is good and accurate the interface could have been better.
Mr. Gayan Jayawardana	The prototype was better than I expected

Table 9.6 - Implementation of the Prototype

9.6.1.4 Context of the Thesis

Question	Can you give your opinion on the content of the thesis?
Evaluator	Feedback
Mrs. Sumithra Pasqual	Overall thesis is good and has much information about the research. Some sections are yet to be filled other than that the content is better.
Mr. Nishantha Perera	The literature review has a better overview of the research as it contains many details about the project the introduction chapter was also very informative. The content of the report is good.
Mr. Nisal Suranaka	Good coverage of existing work and domain in the literature review as it provides what needs expect from the project. Some of the content can be improved overall the content is good, a good research has been done.
Akshina Fernando	The literature review of thesis has a good coverage of existing work and the technologies used in those works. The designs can be improved.
Mr. Arosha Mendis	The thesis has content which are useful for the research.

Mr. Gayan Jayawardana	The content of each chapter in the thesis gives the correct information and useful information was there when I read it. Overall, the thesis was fine.
-----------------------	--

Table 9.7 - Context of the Thesis

9.6.2 Focus Group Testing

9.6.2.1 Prototype Features

Evaluator	Feedback
Tester 1	The system was able to give out the correct answer for the arithmetic problem I gave.
Tester 2	The system features were very well accurate as the answer for the problem was correct.
Tester 3	The core functionality is working well and the answers for the questions that was entered were also correct.

Table 9.8 - Prototype Features Testing

9.6.2.2 Usability

Evaluator	Feedback
Tester 1	The system is very simple and easy to use anyone can understand the system.
Tester 2	Simple and easy to use not complicated at all.
Tester 3	Very interactive UI very easy to use. The help page is available to those who don't know how to use the system.

Table 9.9 - Usability Testing

9.7 Limitations of Evaluation

The author had to face difficulties in the evaluation phase. As of the current prevailing situation in Sri Lanka during the time period all the expert evaluation was limited to google meets, phone calls and emails. Also, finding more experts to evaluate the system was hard as many of them were busy.

9.8 Evaluation on Functional Requirements

The evaluation of the functional requirements which were stated in the introduction chapter are evaluated below.

ID	Requirement	Description	Priority Level	Evaluation
FR1	Should generate answers	For the entered Sinhala arithmetic question the system should generate the answer	Must have	Implemented
FR2	Identify Keywords	The system should be able to identify keywords from the entered question	Must have	Implemented
FR3	Preprocess data	Remove unwanted characters from the question to execute the logic	Must have	Implemented
FR4	Question classification	Executing the model and predicting the category of the question	Should have	Implemented
FR5	Interactive UI	A frontend where the user could interact	Should have	Implemented
FR6	Help page	A help page containing the steps on how to use the system	Could have	Implemented

Table 9.10 - Evaluation on Functional Requirements

9.9 Evaluation on Non-Functional Requirements

ID	Requirement	Priority Level	Evaluation
NF1	Accuracy	Should Have	Implemented
NF2	Performance	Should Have	Implemented
NF3	Maintainability	Could Have	Implemented
NF4	Scalability	Could Have	Implemented
NF5	Usability	Could Have	Implemented

Table 9.11 - Evaluation on Non-Functional Requirements

9.10 Chapter Summary

This chapter thoroughly discussed the evaluation methodology and the approach which was followed. Completely different evaluation criteria were defined, compartmentalizing the project into numerous aspects. The evaluation criteria, the choice of evaluators, furthermore as a summarization of the feedback is also presented here.

CHAPTER 10 : CONCLUSION

10.1 Chapter Overview

This chapter provides the conclusions of the project work as measured against the project objectives begun at the start, and therefore the overall success of the research work completed. This includes aims, challenges faced and solutions considered, along with the limitations and future enhancements to the research work. Additionally, this chapter summarizes the expertise acquired through the degree program's modules and how it was applied. Additionally, a contribution to the present research project is presented.

10.2 Achievements of Research Aims & Objectives

10.2.1 Achievement of the Research Aim

“The aim of the research is to create a problem solving system for Sinhala arithmetic problems that consists of multiplication and division.”

The author successfully completed the research work set out to do and achieved the research aim. The end-users, including domain experts evaluated the implemented system both qualitatively and quantitatively. Therefore, as documented, research has been successfully conducted through the design, development, testing, analysis, and evaluation stages.

10.2.2 Achievement of Research Objectives

The research objectives that were identified in the chapter 1 were covered during this research and below is the status of each objective and the evidence.

Objective	Status	Evidence (Chapter)
Problem Identification	Completed	Chapter 1
Literature Review	Completed	Chapter 2
Data Gathering and Analysis	Completed	Chapter 4
Research Design	Completed	Chapter 6
Implementation	Completed	Chapter 7

Testing and Evaluation	Completed	Chapter 8 & Chapter 9
------------------------	-----------	-----------------------

Table 10.1 - Achievement of Research Objectives

10.3 Utilization of Knowledge from the Course

Module	Description
Computer Science Practice	The module helped to get the basic idea on how to work on presentation and report writing.
Programming Principles	The core of the project is based on programming and the much needed knowledge of programming were covered with the two modules of Programming Principles 1 & 2.
Object-Oriented Programming	The module helped to understand design patterns and programming concepts well during the research designing stage.
Software Development Group Project (SDGP)	The kick start for the final project was given by this module as the author and the group worked on a research idea and followed the same concept and methods as the final project module.
Web Design and Development	This module helped author to get the basic understanding of web development.
Advanced Client Side	Web development with the use of libraries was learned in this module and was much helpful for the author to create the frontend of the system.
Mobile User Experience	This module was helpful as it taught on how to create a proper prototype and what concept to follow when creating.

Table 10.2 -Utilization of Knowledge from the Course

10.4 Use of Existing Skills

The academic courses followed by the researcher at IIT were of immense benefit throughout the execution of this research work. However, the author's personal ability in programming and systems analysis which gained from open source contributing and self-learning also played

a key role in the success of this research project. Previous expertise in Python programming languages (learned during programming principles I and SDGP module) supported during implementation.

In particular the author notes the Internship at Daemonlogic Pvt. Ltd, where the use of development tools and basic knowledge about natural language processing was obtained, as well as the knowledge of python was acquired for the development phase of the final year project.

10.5 Use of New Skills

During the time of this research was conducted the author gained many new skills out of all those skill below listed are some important skills that were gained.

- Natural Language Processing – with the help of online resources such as articles and you tube videos were referred to gain the knowledge about NLP
- Question Answering – most of the knowledge was gained during the literature review stage as many research papers and articles were referred for it.

10.6 Achievement of Learning Outcomes

The following set of skills were gained by completing this project.

Description	Learning Outcomes
Learning about the workflow of doing research is one of the top skills authors gained by doing this project. The way a research project is done is completely different from the way which an application project is done. So, this skill can be only gained by doing this kind of project.	LO1, LO2
This project helped the author to improve problem solving and self – learning skills. As the author had to research, do literature review, and take the critical	LO4, LO5, LO6,

design decisions by himself with the help of mentors, doing this project helped a lot in improving these skills.	
Communication with the supervisor was improved. All of the deliverables were shown and received feedback before the final submission.	LO3, LO4

10.7 Problems and Challenges Faced

In this section the difficulties faced during the implementation and how the author has overcome those difficulties are discussed.

Problem Faced	Solution
Lack of dataset	Since the research domain is a less resource area the author had to create the dataset from scratch.
Creating a model without related documentation	The author had to refer to many research articles before the implementation,
Coding through scratch when creating the model	Had to look into many models which are related to the system to get a better idea on how should the model be created also having a novel approach different to the related codes.
Lack of knowledge in subject	The author took the supervisor's advice on how to start the implementation referring to different articles.
Finding evaluators	The author had a hard time find evaluators to evaluate the project

Table 10.3 - Problems and Challenges Faced

10.8 Deviations

The initial project idea solution was to create a CNN base approach where each stage(keyword identification, question identification and mathematical operator identification) implemented with different algorithms, the author went with a simple approach of creating the question classification with a regression and the keyword identification with a simple POS tagging.

Final solution got an accuracy of 94.87% where it was showed a better accuracy than the previous work.

10.9 Limitations of the Research

Following are the limitation of this research.

- MathQ is a Sinhala arithmetic question answer generator hence all the data should be in the language Sinhala.
- The questions should have two numerical values in the question as the system is designed and implemented for question that have two numerical values.
- Only the simple mathematical operations comprised of multiplication, division, subtraction and addition can be archived by this system.
- In order enhance the project such as scaling the dataset the classification dataset should be label properly.

10.10 Future Enhancements

The implementation mostly covered all the requirements in the initial project proposal but nothing is ever perfect. Some of these future enhancements can develop the system to another level.

- The question classification is done using logistic regression overall the accuracy can increase with the use of deep learning mainly with BERT model.
- The system can be optimized to have more than two numerical values in a single question. As the system can only generate answers which have two numerical values.
- The system could be optimized to solving questions that does contain values not in numerical but in Sinhala wording(as an example the numerical value 2 can be replaced with the word “දෙක”)
- The implementation gets the question from the user through an input field, this can be optimized to getting the question using the voice as the user speaks the question verbally and the system captures it and then the system generate the answer. This would be very

helpful as Sinhala typing is difficult but verbally when the question is spoken it is much easier.

- The system can change way of input from the input field to scanning the question from a camera and then generating the answer for it.
- The research has conducted as a supervised learning approach the system could be built with an unsupervised learning approach for better accuracy and bigger scope.
- A feature can be added where to show the steps on how to solve the problem.

10.11 Achievement of the Contribution to Body of Knowledge

- MathQ is a Sinhala arithmetic answer generation where the main target is for students who are studying Sinhala mathematic problems that contains mathematical operation. The main contribution is for the domain of Sinhala Mathematical as the system is solely build for those two domains and anyone who is working on the domain can use it for their benefit.
- Contribution towards the technical side is the question classification is archived using logistic regression. As per the related work on of it used logistic regression as the classifier for the system. The accuracy of the classification is at a 97.65% where it surpass all the previous related work.
- The initial project proposal the research gap was identified that the answer generation for Sinhala arithmetic problem that has multiplication and division has not been archived. As the main goal was to implement the system for those two operators as a new approach was found to archive that and the accuracy was also high author decided to implement the system to other two mathematical operator which are addition and subtraction.
- As the research had a hard time of finding dataset the author created two dataset comprised of Sinhala mathematical questions. One dataset was created with label dataset where the question is label according to the category of the question to train and test the classification model. The other dataset has the question with answer to test the system.

10.12 Concluding Remarks

The author had it's own ups and down during the project, author overcame all the difficulties and was able to complete the project and develop the system. The system was highly recommended by the evaluators as the project is based on a domain where only less research has been done. The author took the task and was able deliver it very well manned way. Also, the author gain much needed knowledge in NLP and machine learning. With this research provides opportunity to others to conduct more research on the domain as well as work on the limitation and enhancements to provide a greater solution.

References

Ozyurt, I., Bandrowski, A. and Grethe, J., 2020. Bio-AnswerFinder: a system to find answers to questions from biomedical texts. [online] Semantic Scholar. Available at: <<https://www.semanticscholar.org/paper/Bio-AnswerFinder%3A-a-system-to-find-answers-to-from-Ozyurt-Bandrowski/440ed79c70c272ba3a2f94a8f013b1836fb8f2dd>> [Accessed 15 September 2021].

Purwarianti, A., Tsuchiya, M. and Nakagawa, S., 2007. A machine learning approach for Indonesian question answering system.. [online] Research Gateway. Available at: <https://www.researchgate.net/publication/221173808_A_machine_learning_approach_for_indonesian_question_answering_system> [Accessed 17 September 2021].

Yin, W., Yu, M., Xiang, B., Zhou, B. and Schütze, H., 2016. Simple Question Answering by Attentive Convolutional Neural Network. [online] ResearchGate. Available at: <https://www.researchgate.net/publication/303921830_Simple_Question_Answering_by_Attentive_Convolutional_Neural_Network> [Accessed 24 October 2021].

Jayakody, T., Gamlath, T., Lasantha, W., Mallawarachchi, Y., Premachandra, K. and Nugaliyadde, A., 2017. “Mahoshadha”, the Sinhala Tagged Corpus Based Question Answering System. [online] Research Gate. Available at: <https://www.researchgate.net/publication/304614961_Mahoshadha_the_Sinhala_Tagged_Corpus_Based_Question_Answering_System> [Accessed 19 September 2021].

Chathurika, M., De Silva, C., Raddella, A., Ekanayake, E., Nugaliyadde, A. and Mallawarachchi, Y., 2016. Solving Sinhala Language Arithmetic Problems using Neural Networks. [online] Available at: <https://www.researchgate.net/publication/305262356_Solving_Sinhala_Language_Arithmetic_Problems_using_Neural_Networks> [Accessed 24 October 2021].

Javad Hosseini, M., Hajishirzi, H., Etzioni, O. and Kushman, N., 2021. Learning to Solve Arithmetic Word Problems with Verb Categorization. [online] Research Gate. Available at: <https://www.researchgate.net/publication/301405087_Learning_to_Solve_Arithmetic_Word_Problems_with_Verb_Categorization> [Accessed 24 October 2021].

Dwivedi, S. and Singh, V., 2013. *Research and Reviews in Question Answering System*. [online] ResearchGate. Available at: <https://www.researchgate.net/publication/259525702_Research_and_Reviews_in_Question_Answering_System> [Accessed 1 June 2022].

Reddy, A. and Madhavi, D., 2017. *A Survey on Types of Question Answering System*. [online] Iosrjournals.org. Available at: <<https://www.iosrjournals.org/iosr-jce/papers/Vol19-issue6/Version-4/D1906041923.pdf>> [Accessed 1 June 2022].

ALTA Language Services. 2019. *5 Fascinating Facts about the Sinhala Language - ALTA Language Services*. [online] Available at: <<https://www.altalang.com/beyond-words/5-fascinating-facts-about-the-sinhala-language/#:~:text=The%20Sinhala%20language%20is%20syllabic,used%20on%20a%20regular%20basis.>> [Accessed 1 June 2022].

O'Connor, J. and Robertson, E., 1997. *History overview*. [online] Maths History. Available at: <https://mathshistory.st-andrews.ac.uk/HistTopics/History_overview/> [Accessed 3 June 2022].

Cai, L., Zhou, S., Yan, X. and Yuan, R., 2019. *A Stacked BiLSTM Neural Network Based on Coattention Mechanism for Question Answering*. [online] ResearchGate. Available at: <https://www.researchgate.net/publication/335332461_A_Stacked_BiLSTM_Neural_Network_Based_on_Coattention_Mechanism_for_Question_Answering> [Accessed 20 October 2021].

Pennington, J., Socher, R. and Manning, C., 2014. Glove: Global Vectors for Word Representation. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. [online] EMNLP. Available at: <https://www.researchgate.net/publication/284576917_Glove_Global_Vectors_for_Word_Representation> [Accessed 15 September 2021].

Jayaweera, A. and Dias, N., 2022. *HIDDEN MARKOV MODEL BASED PART OF SPEECH TAGGER FOR SINHALA LANGUAGE*. [online] Arxiv.org. Available at: <<https://arxiv.org/ftp/arxiv/papers/1407/1407.2989.pdf>> [Accessed 17 September 2021].

Lutkevich, B., Writer, T. and Burns, E., 2021. *What is Natural Language Processing? An Introduction to NLP*. [online] SearchEnterpriseAI. Available at: <[https://www.techtarget.com/searchenterpriseai/definition/natural-language-processing-NLP#:~:text=Natural%20language%20processing%20\(NLP\)%20is,of%20artificial%20intelligence%20\(AI\).>](https://www.techtarget.com/searchenterpriseai/definition/natural-language-processing-NLP#:~:text=Natural%20language%20processing%20(NLP)%20is,of%20artificial%20intelligence%20(AI).>)> [Accessed 25 June 2022].

Chavan, A., 2019. *Introduction to Conditional Random Fields (CRFs)*. [online] AI Time Journal - Artificial Intelligence, Automation, Work and Business. Available at: <<https://www.aitimejournal.com/@akshay.chavan/introduction-to-conditional-random-fields-crfs#:~:text=CRFs%20have%20the%20ability%20to,are%20dependent%20on%20each%20other.>>> [Accessed 25 June 2022].

Macherla, S., 2018. *Text Classification Using Conditional Random Fields*. [online] Analytics Vidhya. Available at: <<https://www.analyticsvidhya.com/blog/2018/08/nlp-guide-conditional-random-fields-text-classification/>> [Accessed 25 June 2022].

Joby, A., 2021. *What Is K-Nearest Neighbor? An ML Algorithm to Classify Data*. [online] Learn.g2.com. Available at: <<https://learn.g2.com/k-nearest-neighbor#:~:text=The%20k%2>>

[Dnearest%20neighbors%20\(KNN,nearest%20to%20it%20belong%20to.>](#) [Accessed 25 June 2022].

Lutkevich, B., 2020. *What is BERT (Language Model) and How Does It Work?*. [online] SearchEnterpriseAI. Available at: <<https://www.techtarget.com/searchenterpriseai/definition/BERT-language-model>> [Accessed 25 June 2022].

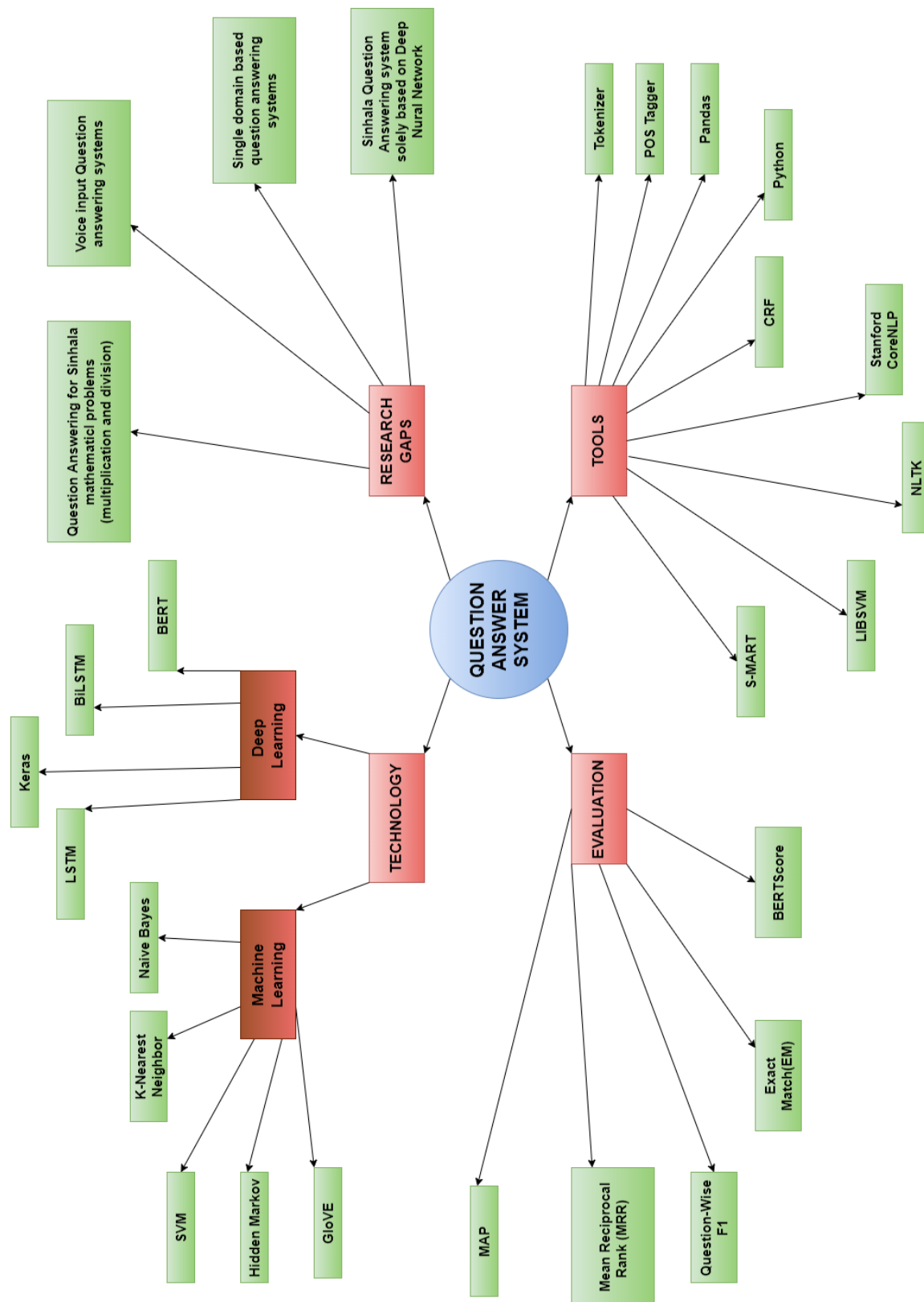
Gandhi, R., 2018. *Support Vector Machine — Introduction to Machine Learning Algorithms*. [online] Medium. Available at: <<https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>> [Accessed 25 June 2022].

Gandhi, R., 2018. *Naive Bayes Classifier*. [online] Medium. Available at: <<https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c>> [Accessed 25 June 2022].

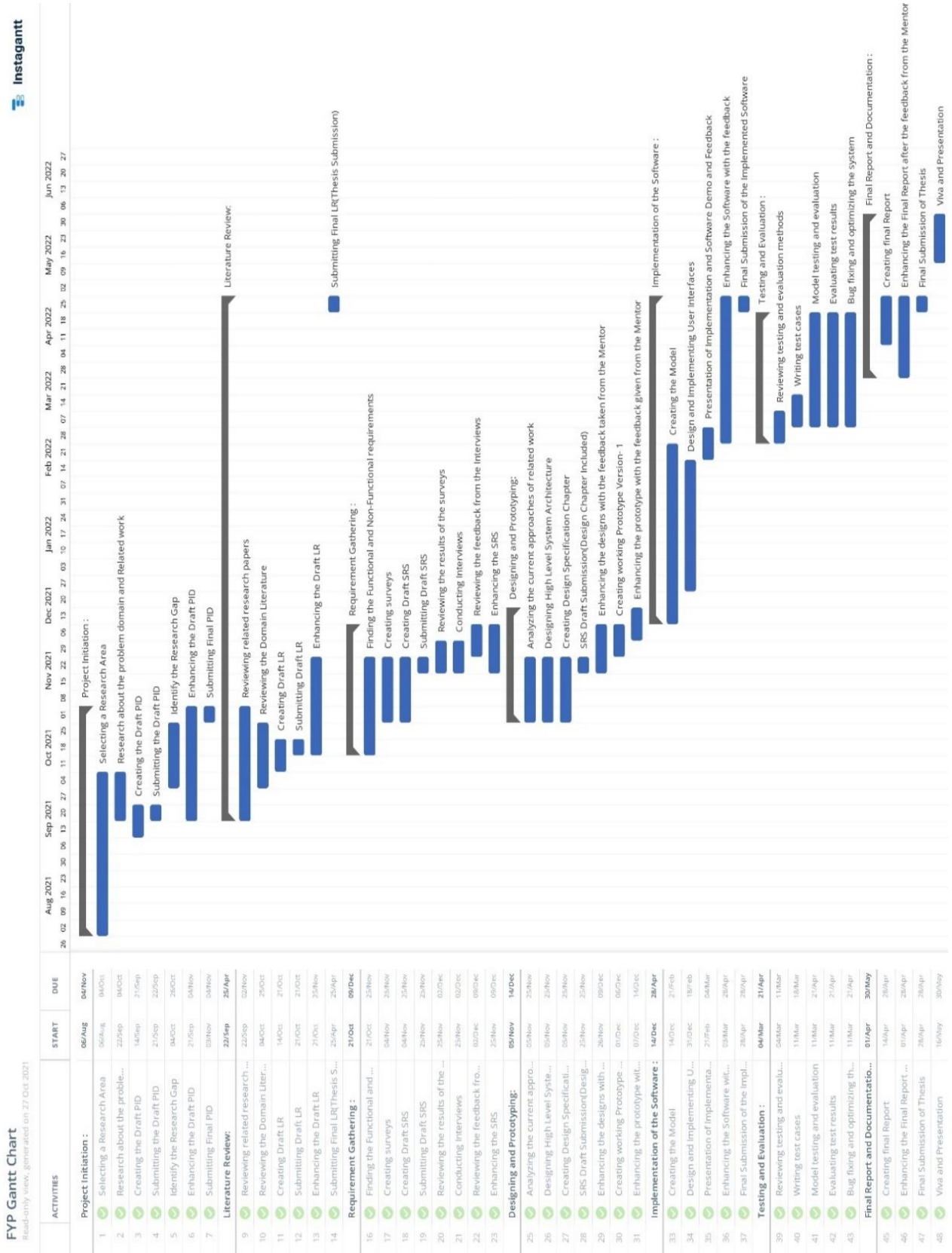
Xu, K., Reddy, S., Feng, Y. and Zhao, D., 2016. *Question Answering on Freebase via Relation Extraction and Textual Evidence*. [online] ResearchGate. Available at: <https://www.researchgate.net/publication/306094046_Question_Answering_on_Freebase_via_Relation_Extraction_and_Textual_Evidence> [Accessed 26 June 2022].

Schubotz, M. and Scharpf, P., 2018. *Introducing MathQA - A Math-Aware Question Answering System*. [online] ResearchGate. Available at: <https://www.researchgate.net/publication/329125386_Introducing_MathQA_-_A_Math-Aware_Question_Answering_System> [Accessed 3 March 2022].

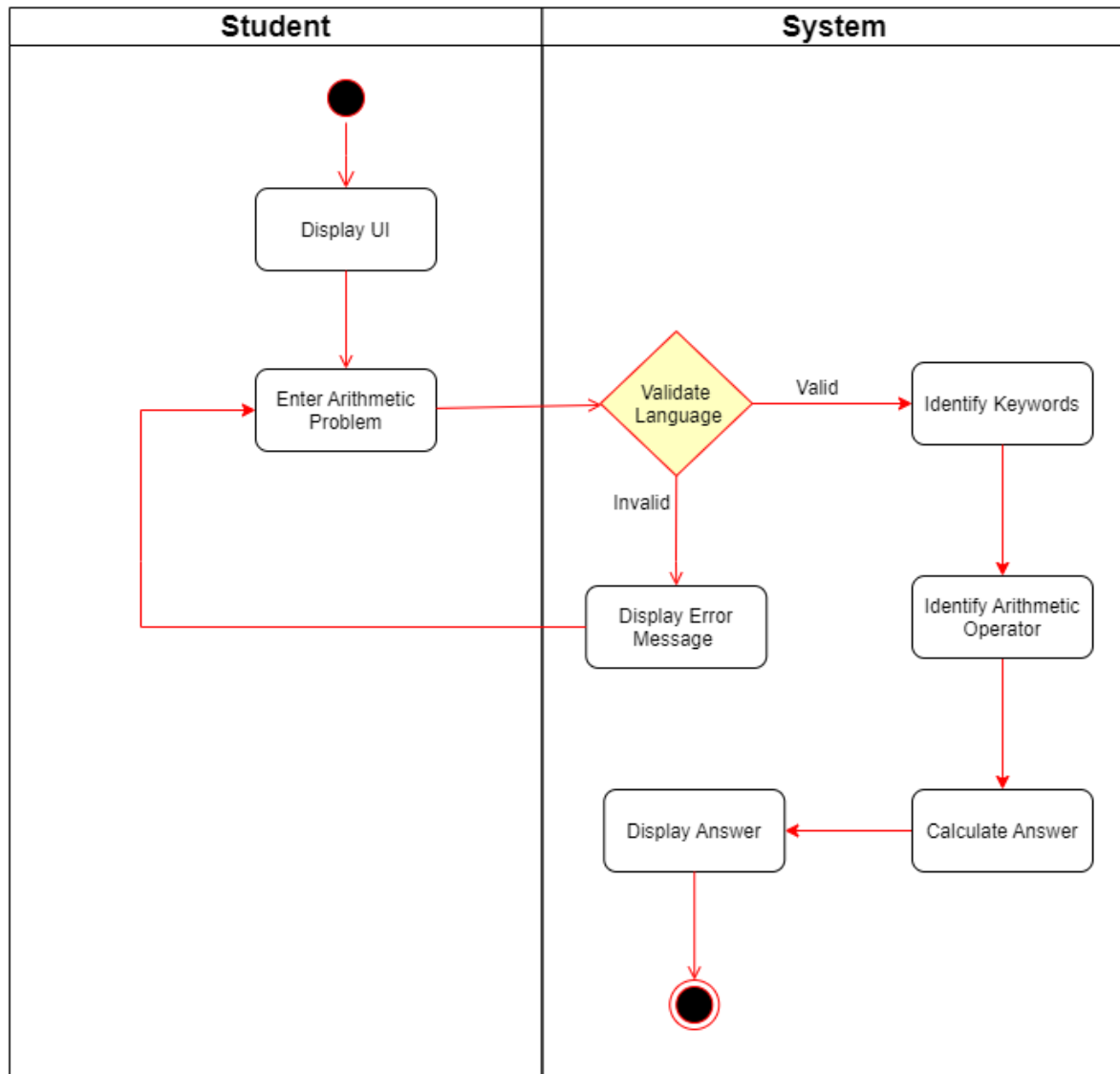
Appendix A – Concept Graph



Appendix B – Gantt Chart



Appendix C – Activity Diagram



Appendix D – UI Designs

Page 1

https://www.example.com

A

Answer Generator

Enter Arithmetic Problem

Generate Answer

Delete Data

Answer

?

Help

Show Steps

Steps

BUTTON

BUTTON

BUTTON

BUTTON

in

f

Appendix E – Requirement Gathering Questionnaire



A Survey on Answer Generator for Sinhala Arithmetic Problems Using Convolutional Neural Network

Hello there,

I am Shenal Anthony, a final year undergraduate student following a BSc (Hons) Computer Science Degree at Informatics Institute of Technology in collaboration with the University of Westminster, UK.

The main target of the research is to create a answer generation system for Sinhala arithmetic problems which contains multiplication and division.

This survey is conducted as a part of my final year research which aims to gather requirements for my final year research project. The survey results will be anonymous and used only for academic purposes.

Please spare a few minutes of your time to fill out this survey.

Thank you very much!

 shenal.2018383@iit.ac.lk (not shared) [Switch account](#)



* Required

Select your age group *

- ☐ Bellow 18
- ☐ 18 - 30
- ☐ 31 - 40
- ☐ Above 40

Are you familiar with Sinhala Arithmetic Problems? *

- ☐ Yes
- ☐ No
- ☐ Maybe

On a scale of 1 to 5 (where 1 been the lowest and 5 been the highest) how would you scale your skills in solving Sinhala Arithmetic Problems? *

- | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1 | 2 | 3 | 4 | 5 |
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

How would you normally solve such problems? (eg: වෙළඳසැලක ඇති පරිපූර්ණ ප්‍රමාණයෙන් $\frac{3}{7}$ ක් ඉතිරිවන සේ පළමු මාසය තුළ විකුණන ලදී. පළමු මාසය අවසානයේදී විකුණා ඇති පරිපූර්ණ ප්‍රමාණය 28kg නම් වෙළඳසැලේ තිබූ මුළු පරිපූර්ණ ප්‍රමාණය කොපමණද?) *

- ☐ Work on your own
- ☐ Google the Answer
- ☐ Ask someone
- ☐ Other: _____

Do you find these questions to be complicated when you try to solve them? *

- ☐ Yes
- ☐ No
- ☐ Maybe

Have you ever encountered system that generate answers for Sinhala Arithmetic Problems? *

☐ Yes

☐ No

If yes, what is the name of that system?

Your answer

Do you think such system would be useful to solve Sinhala Arithmetic Problems? *

☐ Yes

☐ No

☐ Maybe

What features would you like to see in this system?

Your answer

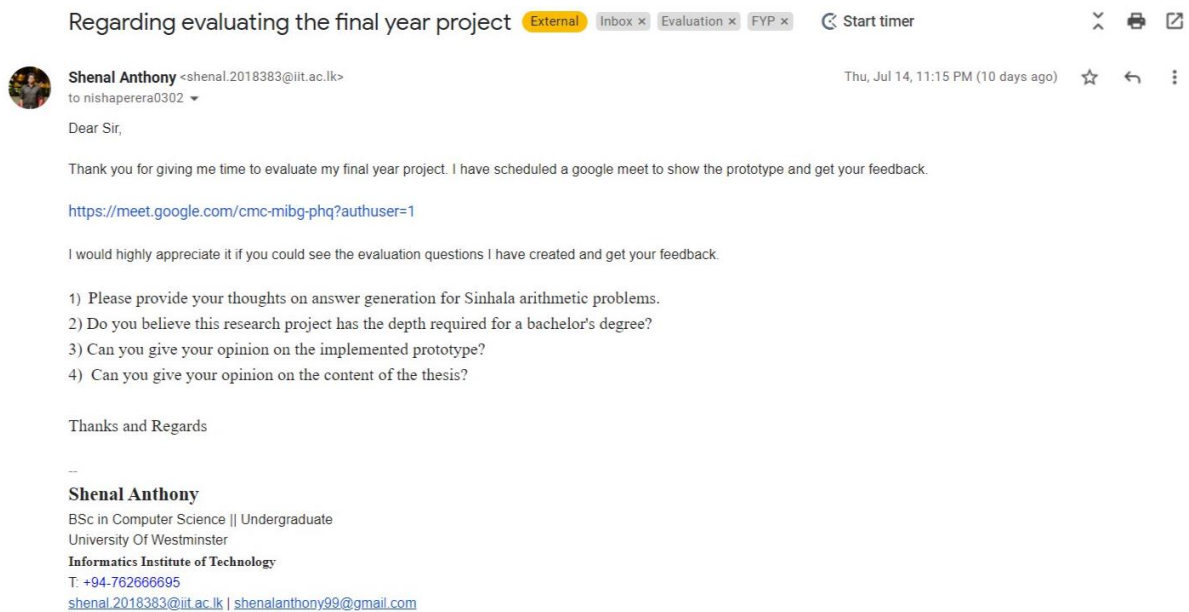
Submit [Clear form](#)

Never submit passwords through Google Forms.

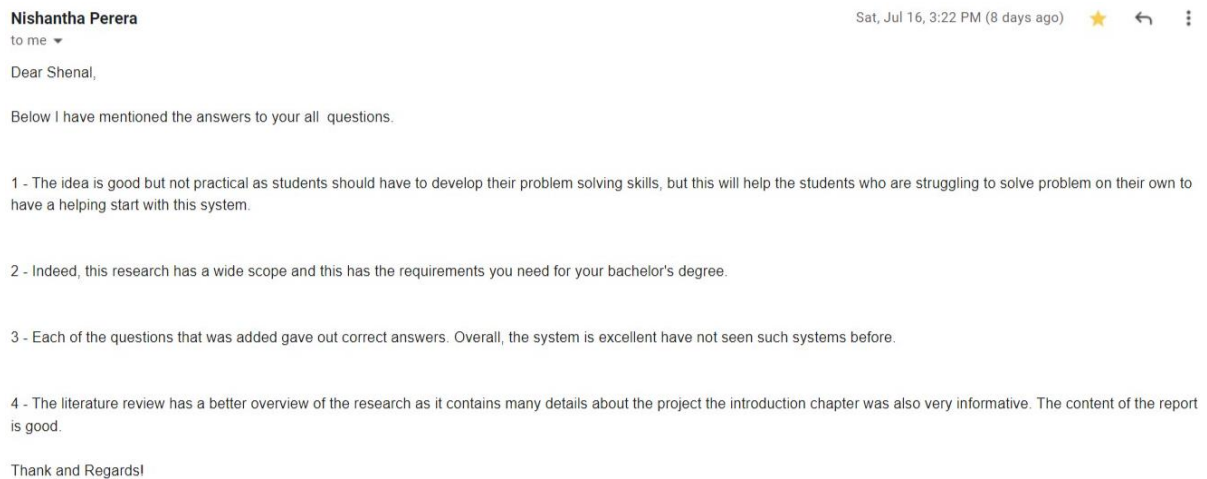
This form was created inside of Informatics Institute of Technology. [Report Abuse](#)

Google Forms

Appendix F – Evaluation Question Email



Appendix G – Evaluation Responses of Expert Evaluators



MathQ - Answer Generator for Sinhala Arithmetic Problems

Sumithra Pasqual

to me ▾

Sat, Jul 16, 11:48 AM (8 days ago) ★ ↶ ⋮

Dear Shenal,

Thank you for selecting me as the evaluator of your final project. Here are the answers to your all questions. I hope it helps to complete your research.

1. The research is a good research area as this would help a lot of students as in the future this would be helping as the students get involved with technology and finding answers for problems they encounter while studying.

2. Yes, The research has a good depth that will cover for your bachelor's degree.

3. The prototype is well implemented the answers are also correct. The system is very accurate on generating answers for mathematical problems.

4. Overall thesis is good and has much information about the research. Some sections are yet to be filled other than that the content is better.

If you have any questions feel free to contact me.

Thank you,
Best Regards,
Sumithra Pasqual

akshina fernando

to me ▾

Sat, Jul 23, 11:17 AM (22 hours ago) ☆ ↶ ⋮

Dear Shenal,

I am very pleased to evaluate your project. Please find my feedback for the respective questions.

1 - The project idea is good research as the domain is Sinhala. Sinhala domain is still an upcoming domain for researchers in the industry as well. Question answering has a wider research as there is much research based on that, creating a question answering system for the Sinhala domain is a better project in my opinion.

2 - Answer generation for Sinhala arithmetic problems have a good depth as the Sinhala domain is a less researched area and it is still a developing area. So, this research in my opinion has depth which is required for a bachelor's degree.

3 - To train and test the model could have used more complex arithmetic problems but as per the scope the implementation is better. The frontend is very user friendly. The basic arithmetic problems are being identified correctly and the answer that has been generated is correct.

4 - The literature review of the thesis has a good coverage of existing work and the technologies used in those works. The designs can be improved.

Wishing you all the very best with your final year project and thesis!

Regards,
Akshina Fernando.

Suranaka, Nisal 341

to me ▾

Thu, Jul 21, 8:13 PM (3 days ago) ☆ ↶ ⋮

Hi Shenal,

Please find my feedback on your final year project:

1 - The whole idea of generating answers for Sinhala arithmetic problem looks easy but as a technical expert I see the depth of this project and the complexity of the project hence the project is a good way to test your skill and develop your knowledge in machine learning/ NLP.

2 - Absolutely, the project contains a good depth which requires for a bachelor's degree in my opinion. As the title might be seen as a project with simple task but in the industry level this has a good value.

3 - A simple UI where any user could use, the implementation has met the project objective as the final output is correct. If the dataset has more records of data then the accuracy can be higher.

4 - Good coverage of existing work and domain in the literature review as it provides what needs expect from the project. Some of the content can be improved overall the content is good, a good research has been done.

Thanks and Best Regards,
Nisal Suranaka | Quality Engineer
Sysco LABS Sri Lanka | 55A Dharmapala Mw, Col. 03
P: +94763317058 | W: syscolabs.lk

Sysco LABS
Technology That Moves Human Civilization

SYSKO'S PURPOSE
CONNECTING THE WORLD TO SHARE FOOD
AND CARE FOR ONE ANOTHER

MathQ - Answer Generator for Sinhala Arithmetic Problems

Arosha Mendis

to me ▼

Sat, Jul 16, 9:51 PM (8 days ago)

Hi Shenal,

The answers for your questions are as follows

- 1 - The project is very helpful as the students would have a great opportunity to solve difficult mathematical problems in the Sinhala language.
- 2 - Yes, I believe this project has the requirements for it.
- 3 - The system is good and accurate, the interface could have been better.
- 4 - The thesis has content which is useful for the research.

Regards,
Arosha Mendis

Gayan Jayawardena

to me ▼

Sat, Jul 16, 11:28 AM (8 days ago)

Hi Shenal,

These are the answers for your evaluation questions.

- 1 - The research is much helpful to society as this would help students solve mathematical problems easily without any hesitation.
- 2 - Yes, the project scope has wide coverage that would cover a bachelor's degree.
- 3 - The prototype was better than I expected
- 4 - The content of each chapter in the thesis gives the correct information and useful information was there when I read it. Overall, the thesis was fine.

regards,
Gayan Jayawardena