Sri Lanka Institute of Information Technology

Arithmetic Problem Solving System for Grade 5 Students using neural networks

(මහෞෂධ2)

Software Requirements Specification (SRS) - CDAP-I

Bachelor of Science (Special Honors) in Information Technology

Contents

Table Of Figures

Tables

Declaration

I hereby declare that this proposal document or part of it is not a copy of a document done by any organization, university, institute or a previous student project group at SLIIT and is not copied from the Internet or other sources except where stated.

Chathurika W.M.T.	IT 12142538	

.....

Mr.Yashas Mallawaarachchi

Supervisor

.....

Mr.Anupiya Nugaliyadde

Co-Supervisor

.....

Dr. Koliya Pulasinghe

Coordinator

1. Introduction

1.1 Purpose

The purpose of this document is to give a detailed description of the requirements for the neural network used in “Mahoshadha 2” system. This is an intelligent system which can solve arithmetic mathematical problems in Sinhala language.

This document will illustrate the purpose and complete declaration for the development of system. It formally specifies the proposed system’s functional requirements, non-functional requirements, data requirements, quality requirements and constraints. The document describes the issues related to the current system and what actions are to be performed by the development team in order to come up with a better solution.

This document will help system developers to understand the overall functionality of the system. By reading this document system users can get a clear idea of systems behavior and how to interact with the system. Also this document can be used to verify whether the software meets the user’s actual needs. The development team uses this document to describe the scope of the project and to plan the system’s design and eventual implementation.

The intended audiences of this System Requirement Specification are, the members of the research group, project supervisor, Mr. Yashas Mallawarachchi, project co – supervisor Mr. Anupiya Nugaliyadda , Project team-

1.2 Scope

This document includes all the details relevant to the neural network which will be implemented for the 1st release of the Mahoshadha 2 system which is planned to be given by December 2015. This neural network is supposed to take the outputs of rest of the 3 functions of Mahoshadha2 system as inputs. The neural network

1.3 Definitions, Acronyms, and Abbreviation

Term	Definition
Neural Network	NN
Artificial Intelligence	AI
Natural Language Processing	NLP
Arithmetic Problem Solving System	APSS

Table 1Definitions,Acronyms,Abbreviation

1.4 Overview

The goal of Mahoshadha 2 system is to enable it's users to take an arithmetic mathematical word problem in Sinhala language , understand it, solve it and simulate how the solution was taken. This system works as an online tutor for students from grade 1 to grade 5.

The neural network that is discussed in this document plays a vital role in the system by ensuring and enhancing the accuracy of the system , and also by generating the final outcome of the system.

2. Overall description

2.1 Product perspective

People always wanted computers to act intelligent. To accomplish this task the field of Artificial Intelligence came into existence. One of the key barriers in making computers intelligent is understanding of Natural Language. Natural language processing which deals with understanding of languages is sub division of Artificial Intelligence. Question Answering is a classic NLP application. It has practical applications in various domains. Most of the available system's Task that a question answering system realizes is given a question and collection of documents, finds the exact answer for the question. In this project we are going to apply some of the Natural Language Processing (NLP) techniques to analyze the Sinhala language to gain a better understanding of the language in an NLP perspective and as a step towards developing more complex tools for algebra word problem solving in Sinhala.

Designing algorithms to solve math problems is a challenge. Arithmetic world problems describing world state and pose question about it so that state can be modeled using equations whose question's answers. For a child language understanding part is little bit tough when we give arithmetic word problem they should identify what are the category they must consider and solve the problem. Arithmetic word problems are an important practical application of mathematics since real world problems usually do not arise in terms of equations but as verbal or pictorial representations.

In our literature survey we found out that the most relevant and best researches at IBM built unstructured information management architecture, a frame work to integrate various computer software each different and independent to work together .It was a real time and scalable question answering system.IBM Watson does not cover up solving mathematical problems. But the project that is to be developed is based on mathematical sentence problem solving. In comparison to

IBM Watson main difference is that “Mahoshada 2“is developed for Sinhala language.

Although there are many advanced related systems for English language, there is no any system for Sinhala language. This research will be a great opportunity to Sinhalese who are expecting this type of system in their native language. This document will further describe our objectives of doing this type of research and how we expect to do the research and many other related details. This is an encouraging start for further research on the Sinhala language.

2.1.1 System Interfaces

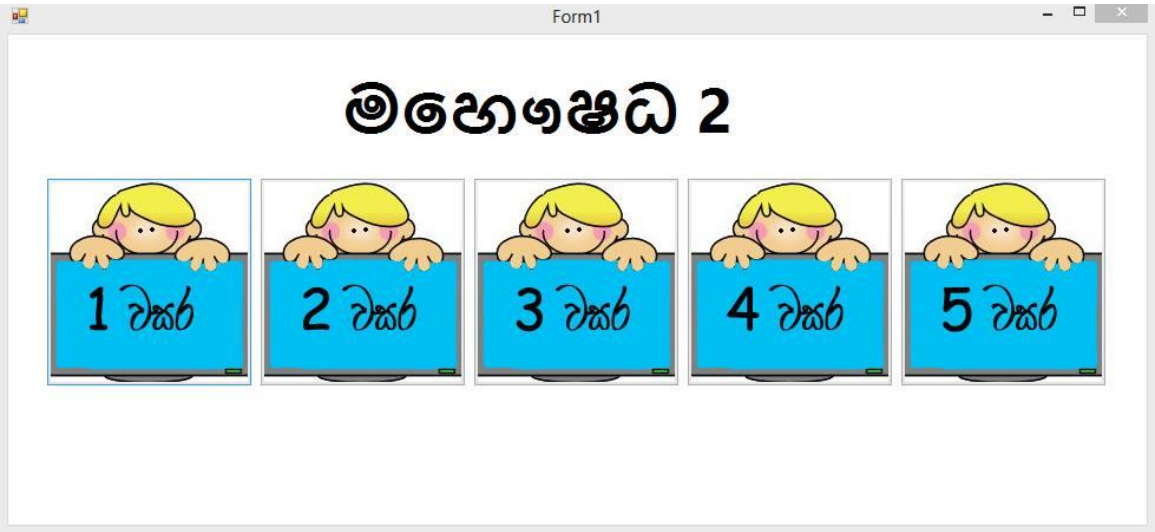

Figure 1 : interface 1

Figure 2 : interface 2

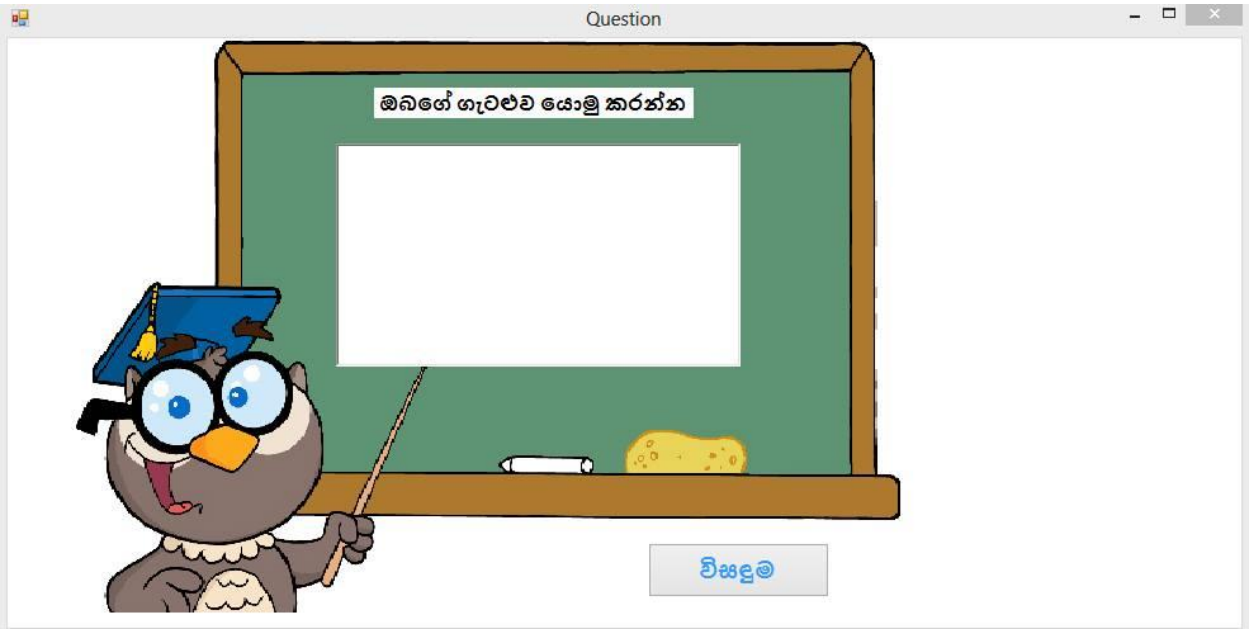

Figure 3 : interface 3

2.2Product functions

- Identifying key words.
- Identifying the question.
- Generating the equation for a problem.
- Producing the final answer for a given problem.
- Checking the answers given by the user.
- Handle inputs through hidden layers.
- Passing output of a hidden layer to another layer.

2.2.1 Use case Scenarios

2.2.1.1 Taking inputs for a hidden layer

Use Case name	Handle inputs in hidden layers
Actor	User , system
Pre-Condition	System is working, outputs have been generated by other functions
Success Scenario	1. Use case scenario starts when the neural network receives keywords to the first hidden layer. 2. Weights are assigned. 3. Sigmoid activation function is applied to weights. 4. Threshold the results. 5. Get the values that exceed the threshold. 6. Outputs are sent to next hidden layer.
Exceptions	5a. if no input value pass through the threshold, layer takes next set of inputs
Post-Conditions	Output must be sent to the next layer.

Table 2Usecase scenario:Taking inputs for a hidden layer

2.2.1.2 Insert a new question to the database

Use Case name	Insert a mathematical problem to the database
Actor	User , system database
Pre-Condition	System is working properly
Success Scenario	1. Use case scenario starts when the user press “add new question” button 2. User is given a text area to insert new question. 3. User enters the word problem. 4. User press enter button.
Exceptions	-
Post-Conditions	Update the database.

Table 3: usecase scenario: Insert a new question to the database

2.2.1.3 Key word extraction.

Use Case name	Key word extraction
Actor	User , system
Pre-Condition	System is running, arithmetic question is received from the user
Success Scenario	1. The use case starts when the system received a arithmetic question 2. Token extraction from given question. 3. Extracting nouns, verbs, adjectives …. etc. 4. Calculate the properties of each candidate. 5. Select possible keywords.
Exceptions	4a. if system get different values for the calculation get the highest value among them. 4b. if system get negative values for the calculation omit that words.
Post-Conditions	System has successfully extracted required keywords

Table 4 : usecase scenario : key word extraction

2.2.1.4 Identifying Mathematical Operation and Generate the equation

Use Case name	Identifying Mathematical Operation and Generate the equation
Actor	User , system
Pre-Condition	System is running, key words are extract from the question, Question identification is done
Success Scenario	The use case starts when the system received extracted key words and question identification is done. 2. Forming the states. 3. Match the entities and categories. 4. Initializes or updates the values of the containers in the state. 5. Quantities of entities update or observe. 6. Using the learned verb categories find the relevant mathematical operation. 7. Forms an equation by comparing the quantities for containers matched between the two states.
Exceptions	3a. An entity/category is matched if it has the same head word and same set of attributes as an existing entity/category. 3b. If an entity or category cannot be matching, then a new state is created. 4a. For the matched entities, initializes or updates the values of the containers in the state.
Post-Conditions	

Table 5 : usecase scenario :2.2.1.4 Identifying Mathematical Operation and Generate the equation

2.2.1.5 Question Identification

Use Case name	Question Identification
Actor	User , system
Pre-Condition	System is running, arithmetic question is received from the user
Success Scenario	1. The use case starts when the system received an arithmetic question 2. Word chunking for the received question 3. Tagging the words 4. Identifying the question words(what, howmany,howmuch etc..) 5. Identify the question
Exceptions	
Post-Conditions	System has successfully identified the question

Figure 4 : Use case scenario : Question Identification

2.3 User Characteristics

The system is to be used by Students from Grade 1 to grade 5. Students are required to have computer literacy to a certain extent. User can either type the mathematical problem on the prompt or copy paste it from somewhere. So sometimes parents can also be involved with the system if the student is not able to type words in Sinhala.

2.4 Constraints

The section below briefly addresses the some of the constraints of the system.

The end user of the system are most of the time children. Even though the technology has being developed and most of the schools are focused on improving the computer literacy of students, there may be students who are less aware of using a computer. So they will need to have the assistance of parents.

On the other hand most of the time the end user does not have a knowledge about the background process.

If we talk about the constraints on the neural network of the system , there are certain constraints.

Neural network is operating on Sinhala words.

3. Neural network structure

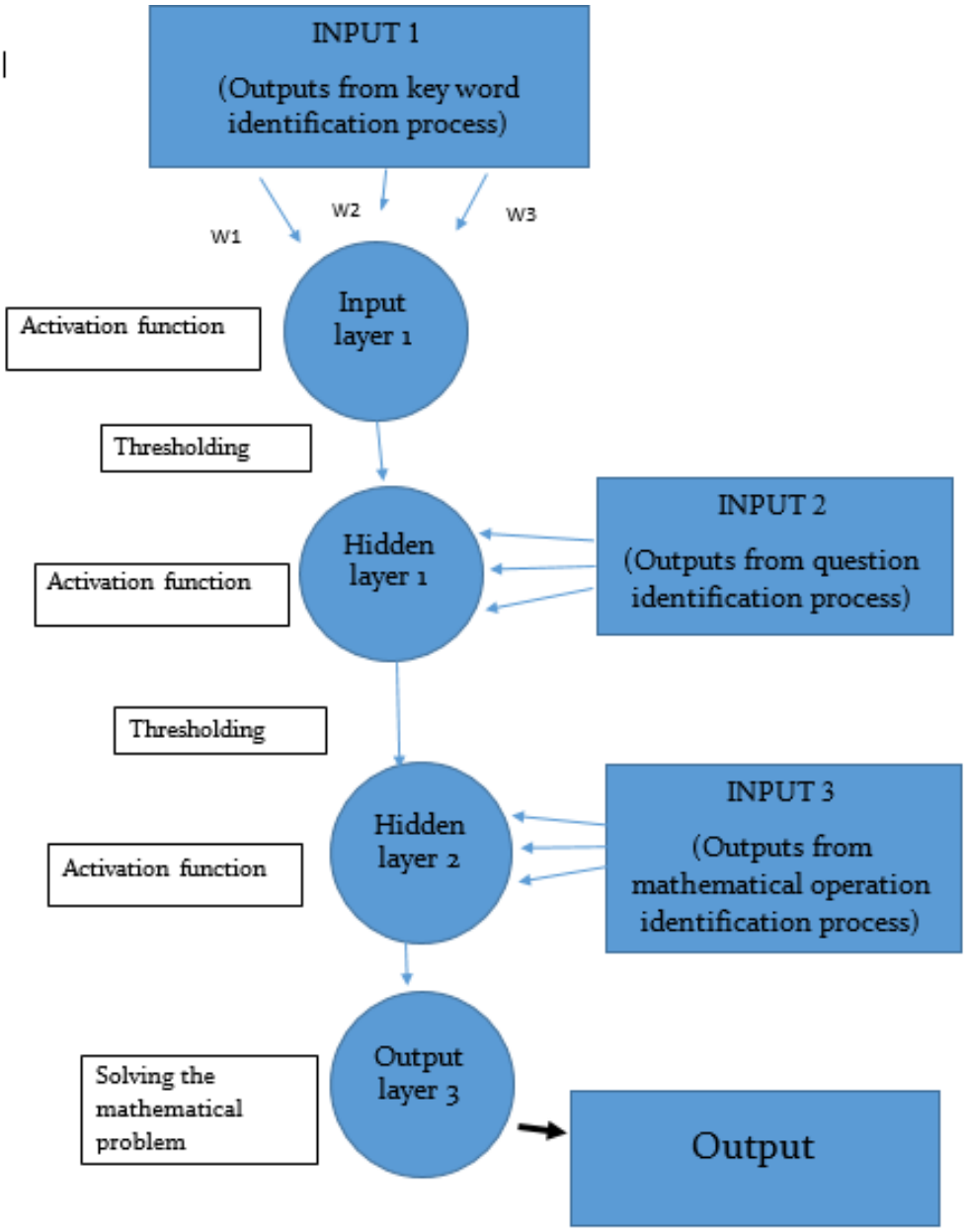

Figure 5 : Neural network structure

As shown below diagram there are 3 inputs for this neural network

- Outputs from keyword identification process
- Outputs from mathematical question identification process
- Outputs from mathematical operation identification process

Inside each hidden layer there will be an activation function that will generate the output of that layer. Most relevant components to generate the solution of the problem will be extracted through this activation function. The outputs that are generated by the activation function will have to be threshold. Thresholding will convert the answers of activation function into 1 or 0.

Only the outputs which hold '1' will be sent to the next layer.

$$\text{output} = \begin{cases} 0 & \text{if } \sum_j w_j x_j \leq \text{ threshold} \\ 1 & \text{if } \sum_j w_j x_j > \text{ threshold} \end{cases}$$

4. System Requirements

4.1 Nonfunctional Requirements

4.1.1 Performance Requirements

The only way in which systems will meet their performance targets is for them to be specified clearly and unambiguously. It is a simple fact that if performance is not a stated criterion of the system requirements then the system designers will generally not consider performance issues. The performance of the product must be to the highest level, since lot of mathematical questions are handled through the system. Since the system consists of lot sub modules coming up with a suitable design specifically concentrating on performance wise is very important.

4.1.2 Safety Requirements

Safety consistent with mission requirements, is designed into the software in a timely, cost effective manner. The system will always be confirmed to run properly and give the correct answer effectively, therefore there is no risk of any data losses. Further data retrieval is also efficient.

4.1.3 Security Requirements

Security is often an afterthought during software development. Realizing security early, especially in the requirement phase, is important so that security problems can be tackled early enough before going further in the process and avoid rework. According to the user privileges of the system, user can perform his /her relevant tasks which are assigned to them by the system.

4.2 Software Requirements and Hardware Requirements

- Server Computer - core i5 2.6 GHz
- Ram - 4 GB RAM
- Hard Drive - 500 GB hard disk drive
- VGA - 1GB
- Linux fedora
- VMware Workstation
- Adobe Photoshop
- MS project
- MS Office