Arithmetic Problem Solving System for Grade 5 Students using Neural Networks

Project team

Supervisor : Mr. Yashas Mallawarachchi
Co Supervisor : Mr. Anupiya Nugaliyadda

- Chathurika W.M.T IT12142538
- Raddella A.M IT12082148
- Ekanayake E.M.R.S IT12142774
- De Silva K.C.E IT12143214

Introduction

- Natural language processing(NLP) deals with understanding of languages .
- Question Answering(QA) is a classic NLP application that has practical applications in various domains.
- Most of the available QA systems' task is extracting answer from documents.
- Most of these systems are not working on solving Arithmetic word problems in Sinhala language.

IT12143214

Current systems

- **IBM Watson**
- **Start natural language question answering system**

  Refer documents and extract the answers.

START, the world's first Web-based question answering system, has been on-line and continuously operating since December, 1993. It has been developed by Boris Katz and his associates of the InfoLab Group at the MIT Computer Science and Artificial Intelligence Laboratory. Unlike information retrieval systems (e.g., search engines), START aims to supply users with "just the right information," instead of merely providing a list of hits. Currently, the system can answer millions of English questions about places (e.g., cities, countries, lakes, coordinates, weather, maps, demographics, political and economic systems), movies (e.g., titles, actors, directors), people (e.g., birth dates, biographies), dictionary definitions, and much, much more. Below is a list of some of the things START knows about, with example questions. You can type your question above or select from the following examples. less...

Geography

- What South-American country has the largest population?
- What's the largest city in Florida?
- Give me the states that border Colorado.
- What cities are within 250 miles of the capital of Italy?
- How many people live in Israel?
- Show me a map of Denmark.
- Which is deeper, the Baltic Sea or the North Sea?
- How far is Mount Kilimanjaro from Mount Everest?
- List some large cities in Argentina.

Science and Reference

- What is Jupiter's atmosphere made of?
- Who first discovered radiocarbon dating?
- How far is Neptune from the sun?
- Why is the sky blue?
- What planet has the smallest surface area?
- How many feet are there in a kilometer?
- Convert 100 dollars into Euros.
- Show me a metro map of Moscow.
- How many languages are spoken in Afghanistan?

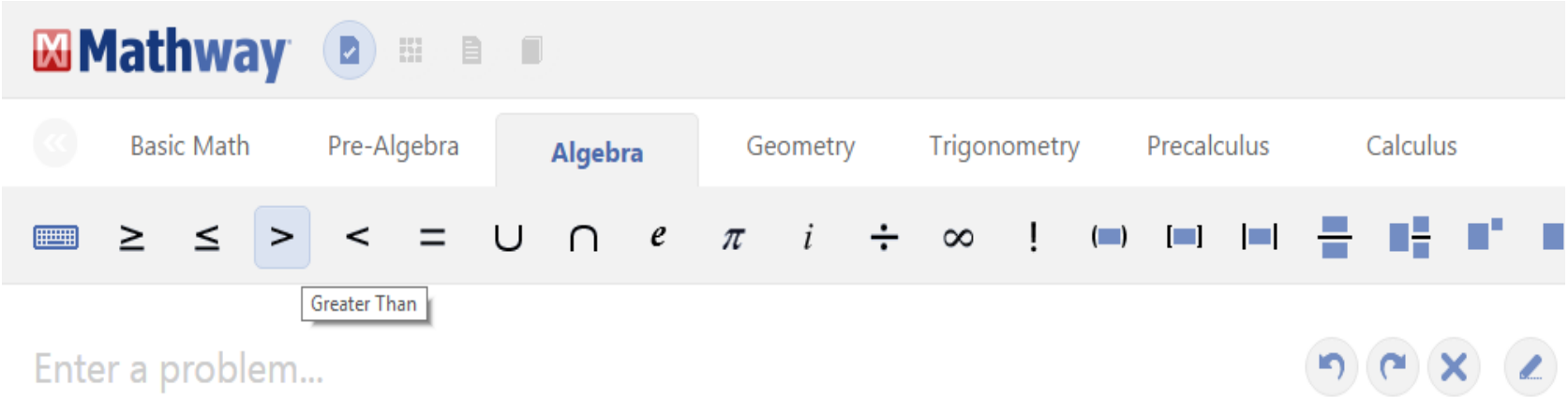
Current systems

- **Aris**

  Only capable of understanding English language

- **mathway.com**

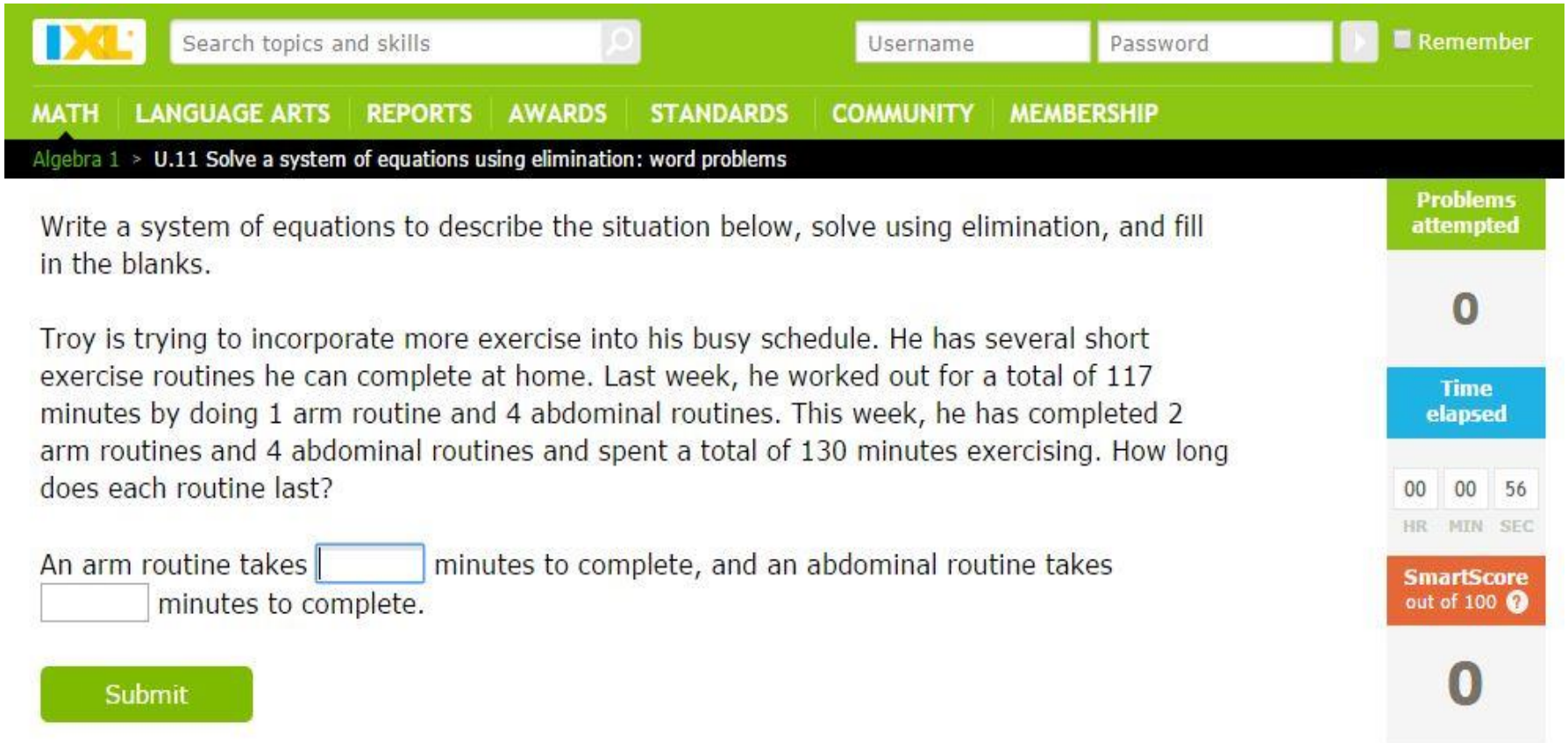
  only solving direct mathematical questions



IT12143214

Current systems

- IXL



IT12143214

Differ From Existing Systems

- Related systems for question answering are in English language.
- proposed system's main target is solving Arithmetic word problems in Sinhala language.
- Most of them are available for information retrieval only.
- Capable of understanding a detailed sentences
- System can solve any given problem by itself

Research Problem

- Come up with a solution to understand the Sinhala language of the system .
- Identify basic math facts which is very important for problem solving (keywords) .
- Determine which mathematical approach to use to solve a problem (mathematical operation)

IT12082148

General objectives

Develop a system that is capable of solving grade 5 mathematical word problems.

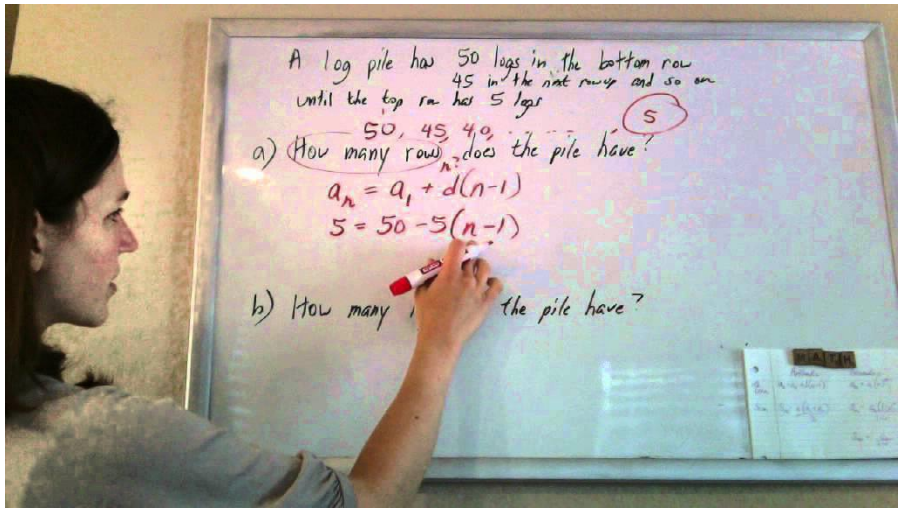
Take word problem as text → Understand the question → Understand the Required operation

SOLUTION

IT12082148

Specific objectives

- Maintain better interaction with user and the system.
  - Provide an attractive user interface that suits for students in age of 9 – 10.
  - Providing a user friendly interface for users



IT12082148

Specific objectives

- Cover up all the areas in the scholarship syllabus that deal with addition and subtraction operations

Addition in number theory

Subtraction in number theory

Currency

Fractions

Time calculations

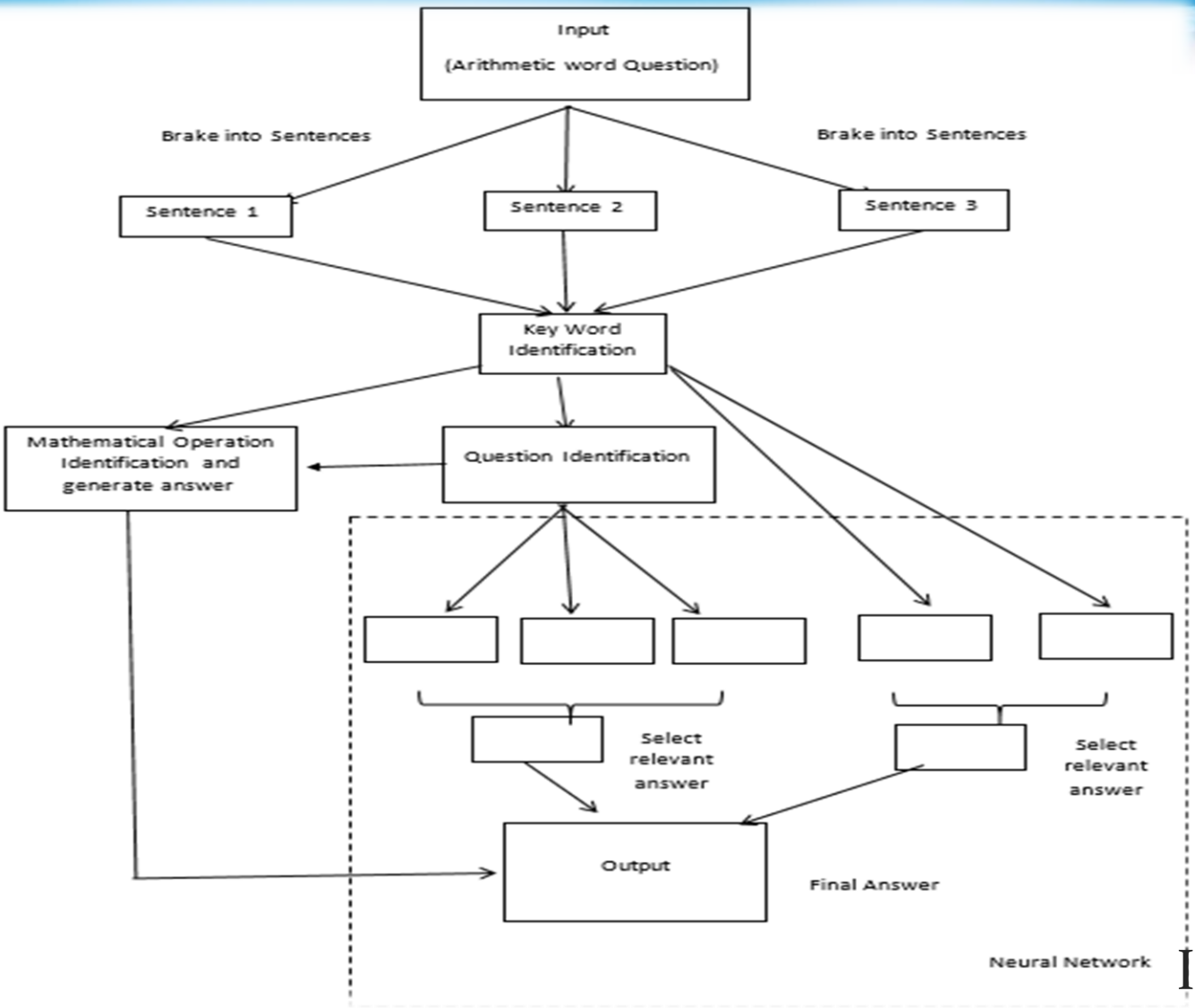
Decimal calculations

IT12082148

Specific objectives

- Understand The Correct mathematical problem
- Make the final output more understandable for the user
- Increase the accuracy of the system

System diagram



IT12082148

System is divided into 4 sub modules.

Member1
- Identifying Keywords

Member 2
- Mathematical question identification

Member 3
- Identifying Mathematical Operation and generating an equation

Member 4
- Neural network

IT12142774

Methodology

Identifying Keywords

Candidates :

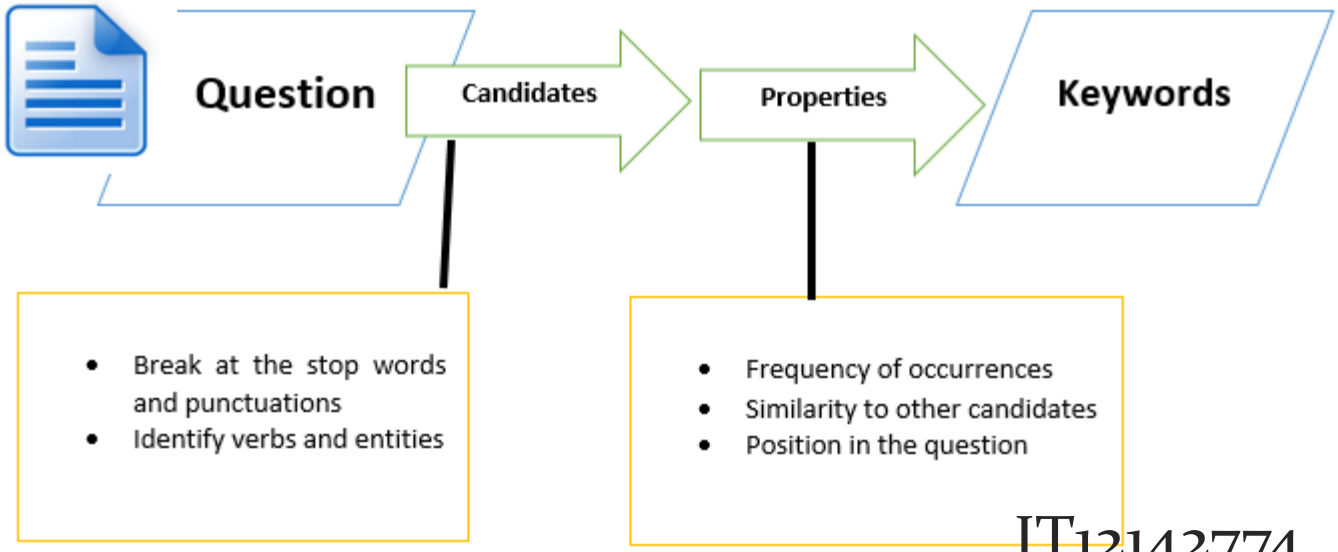
- Use morphological parser analyze and generate Sinhala verbs.

- Use the named entity recognition system to identify the entities (numbers and people etc..).

Methodology

Properties :

- Use PPMI (Positive Point wise Mutual Information) to identify the word similarity.
- To identify more suitable keywords develop a new PMI metric.



Methodology

Mathematical question identification

මාලා ලග කළු පුසන් 10ක් සිටි.ඇය කළු පුසන් කීපදෙනෙක් සමන්ට දුන්පසු මාලා ලග කළු පුසන් 4ක් ඉතිරිව සිටියි. සමන්ට ලැබුනු කළු පුසන් ගණන කියද.?

- To identify question, have to create a feature function f_i

In a Conditional Random Fields (CRF), each **feature function** is a function that takes in inputs as:

- a sentence s
- the position i of a word in the sentence
- the label l_i of the current word
- the label l_{i-1} of the previous word

IT12142774

Methodology

Assign each feature function f_j a **weight** λ_j

- Given a sentence s, we can now score a labeling l of s by adding up the weighted features over all words in the sentence

$$\text{Score } (l|s) = \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_i \, f_i \; (s, i, l_i, l_{i-1})$$

Methodology

- Finally, we can transform these scores into probabilities $p(l|s)$ between 0 and 1 by normalizing.

$$P(l|s)=\frac{\exp[score(l|s)]}{\sum l^1 \exp[score(l^1|s)]} = \frac{\exp[\sum_{j=1}^{m}\sum_{i=1}^{n} \lambda_j f_j(s,i,l_i,l_{i-1})]}{\sum_{l^1} \exp[\sum_{j=1}^{m}\sum_{i=1}^{n} \lambda_j f_i(s,i,l_i^1\ l_{i-1}^1)]}$$

- $f_1(s,i,l_i,l_{i-1}) = 1$ if l_i=question word , and the sentence ends in a question mark; 0 otherwise.

- Again, if the weight λ1 is large and positive, then labeling that assign VERB to the first word in a question (e.g., "Is this a sentence beginning with a verb?") are preferred.

IT12142774

Methodology

Identifying Mathematical Operation and generating an equation

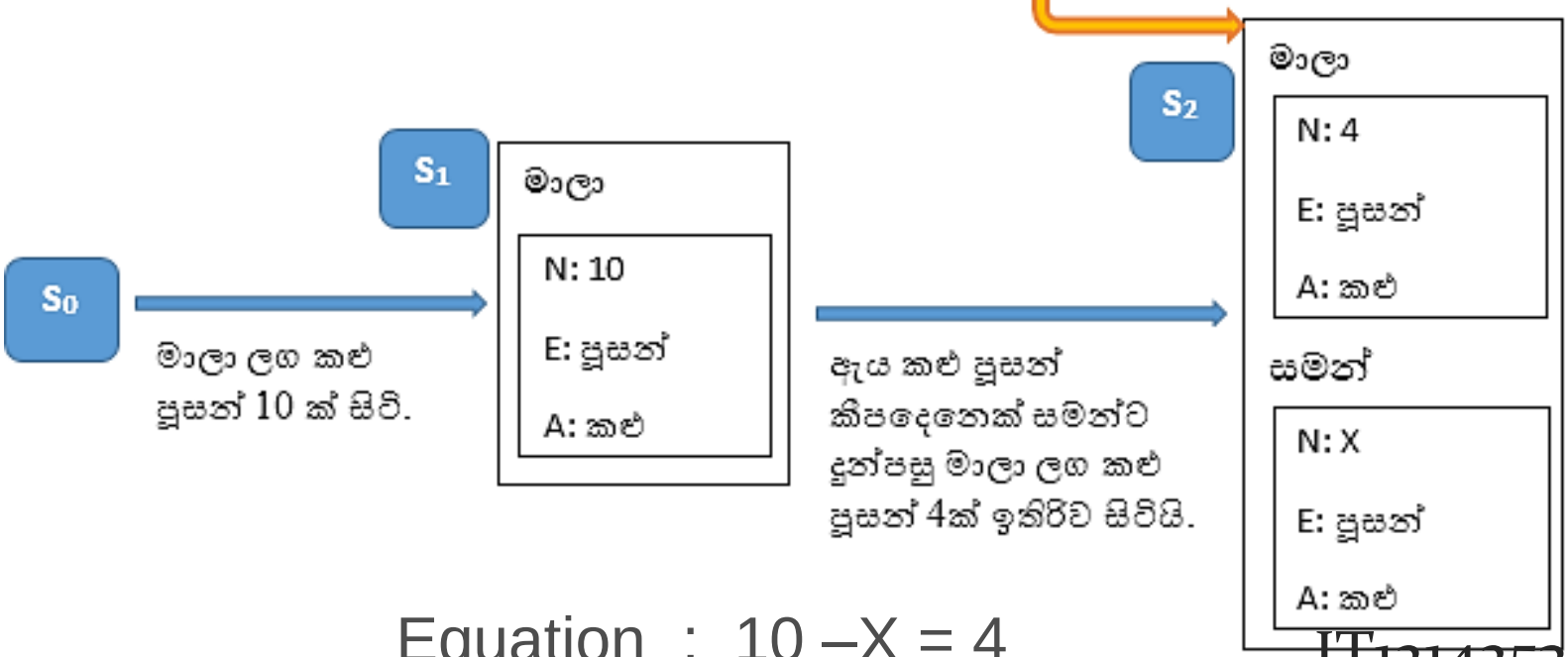
1) Progressing states based on verb categories in sentences.
    - After finding entities and quantities, form the states.
    - Match the entities and categories .
    - For the matched entities, initializes or updates the values of the containers in the state using verb categories.

Methodology

02)Forming the equation.

Question 01

මාලා ලග කළු පූසන් 10 ක් සිටී. ඇය කළු පූසන් කීපදෙනෙක් සමන්ට දුන්පසු මාලා ලග කළු පූසන් 4ක් ඉතිරිව සිටියි. **සමන්ට ලැබුන කළු පූසන් ගණන කීයද?**



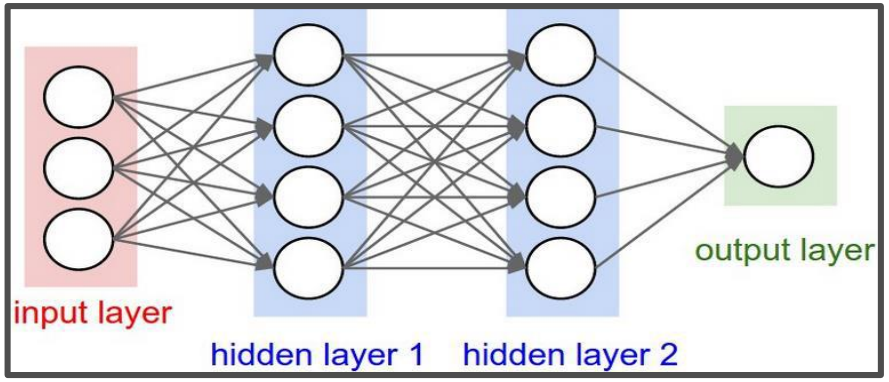
Equation : 10 –X = 4

IT12142538

Methodology

Neural network

- Two hidden layer neural network including input and output layers.



- Feed forward system

- Sigmoid activation function

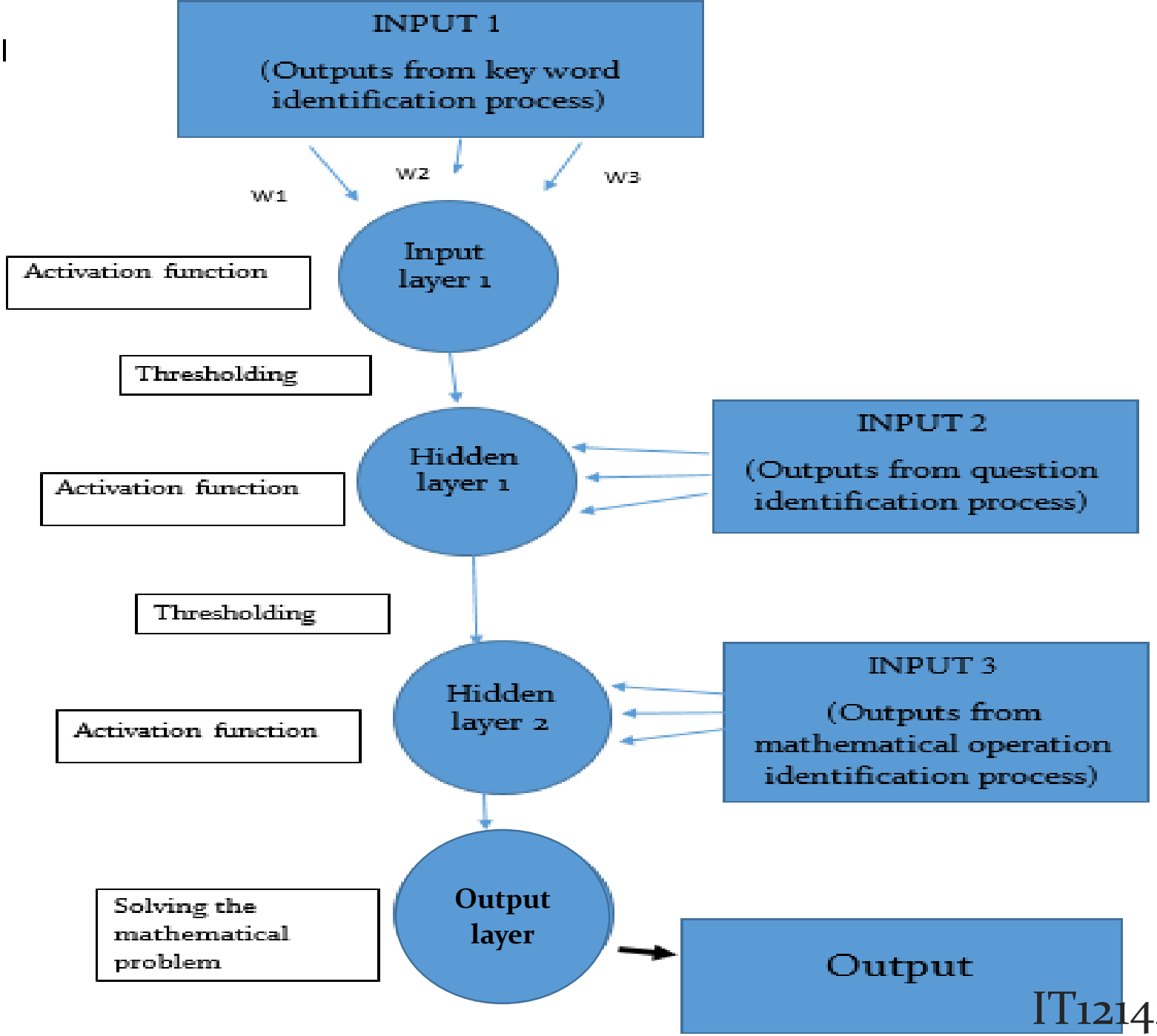
$$sig(t) = \frac{1}{1 + e^{-t}}$$

IT12142538

Methodology

Neural network cont..

- Train the network through supervised learning.
- 'Online training' will be performed in training the network


INPUT 1
(Outputs from key word identification process)
W1
W2
W3
Activation function
Input layer 1
Thresholding
Hidden layer 1
INPUT 2
(Outputs from question identification process)
Activation function
Thresholding
Hidden layer 2
INPUT 3
(Outputs from mathematical operation identification process)
Activation function
Solving the mathematical problem
Output layer
Output

IT12142538

Users and benefits

Users

- Students in grade 1,2,3,4,5

(Future Work)
students up to grade 11

Users and benefits

Benefits

- Students can learn mathematical problem solving in Sinhala Language.
- Students can learn how to solve the question through simulations
- Students can learn any mathematical problem being in the boundary of topics that the system covers.

Marketing plan

- Around 300000 students face grade 5 scholarship examination annually.
- Computer literacy of students is increasing day by day.
- Most of the schools are having computer laboratories now a days.
- Students are more interested in using new technology more than books.
- Introduce the free versions of system for schools and encourage students to use the system.

IT12142538

THANK YOU