**Deep Neural Network based system for solving Arithmetic Word Problems**

**Basic Approach** --> The goal is to convert arithmetic problem to mathematical equation using NLP.

DILTON named model is used that can only operate upon (+,-,/,\*) operators.

The main concept is to divide the arithmetic problem into two parts i.e. worldstate and query. Worldstate can be defined as everything in the question that contains numerical and all relevant/irrelevant information whereas query can be regarded as the real question on which the type of algorithm we will be using will be depending.

e.g. I had 2 mangoes last week and this week I have 4 mangoes. How many mangoes did I got this week?

Worldstate --> I had 2 mangoes last week and this week I have 4 mangoes.

Query --> How many mangoes did I got this week?

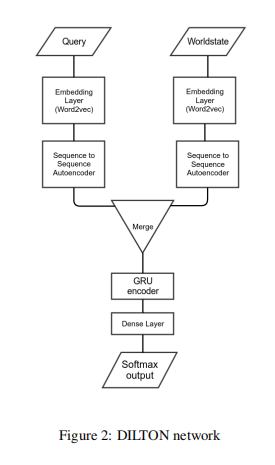
The query part will help us to understand the operation we’ll be doing upon question like in this example we know we need to subtract 4 from 2 and that was depicted by query part.

In symbolic approach, mathematical problem sentence are transformed by pattern matching or verb categorization.

The model uses GRUs and LSTMs to process question and predict the operation needed to be operated upon them.

**Model Architecture** –-> The pipeline consists of three different type of models: -

1. Sequence Autoencoder – It will be operated upon both worldstate and query to process the problem sentence, word2vec embeddings is used.
2. Combining the representation – The output from both worldstate and query will be combined together to produce final output.
3. Predicting the answer – After combination of both answers, to get final output GRUs are used to get the final vectorized version of answer which can be decoded to get the proper output.



**Training** –> In order to train our Neural Network, 30% Dropouts is reserved and 50 word2vec embeddings are taken in account with 100 hidden nodes of GRU both for worldstate and query, the model is trained for 40 epochs and provide 99% training accuracy and 88.8% testing accuracy.

**Dataset used** –> The data set used is provided by MAWPS and consist of single operation questions, training set consist of 1314 questions and testing set consist of 438 questions.

**Problems Associated** –> The model can’t be implemented on equations having more than two operands or more than one operator.

**World state(vectorize)**

**Query(vectorize)**

**Stop word removed world state and query**

**output**