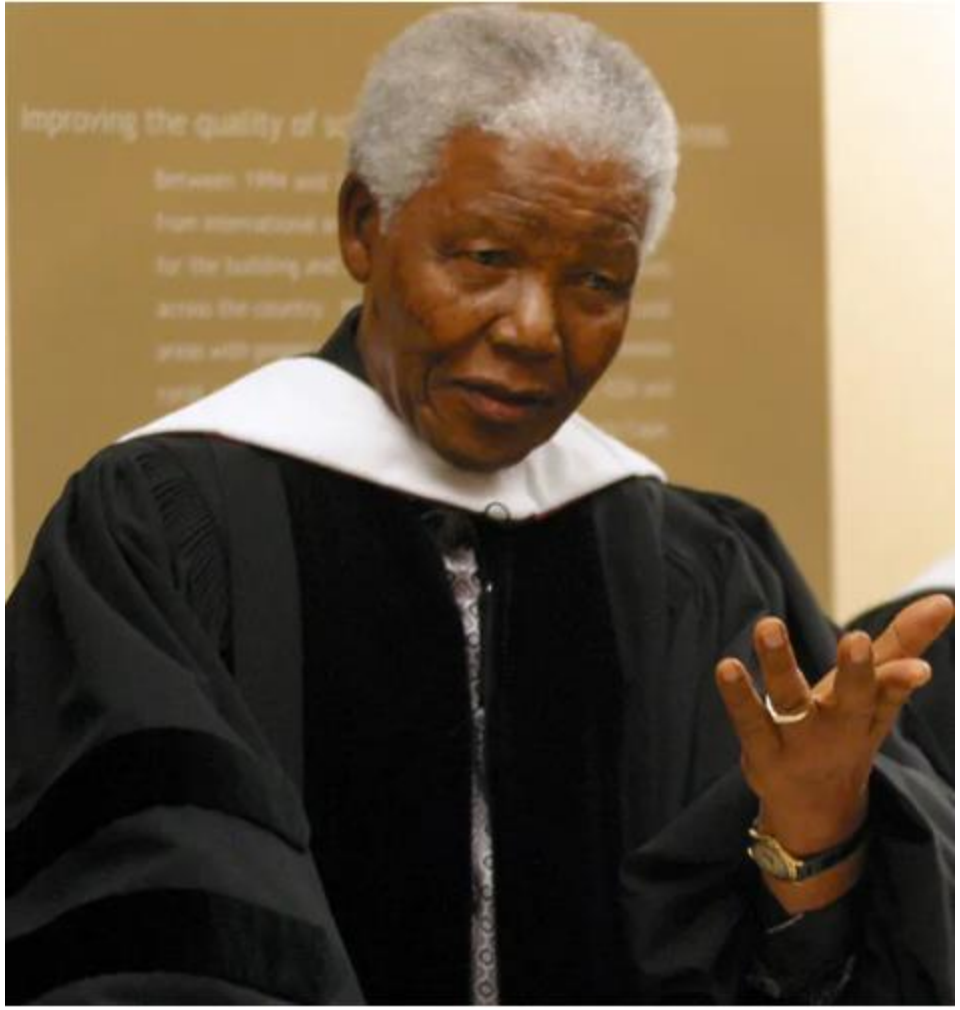
Data Science Africa 2019, Accra, Ghana
Ashesi University

# Subjective Question Marking using Deep Learning

**Abebaw Eshetu**

**Haramaya University, Ethiopia**



*"… Achievement of greater social justice is closely dependent on equitable access by all sections of the population to quality education."*

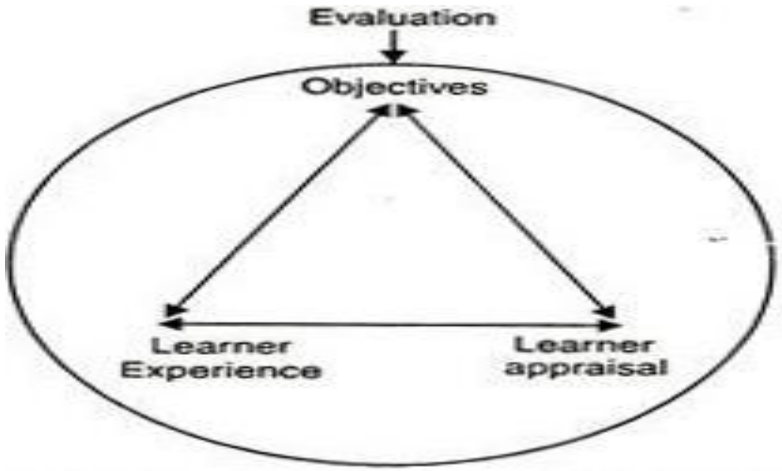
***"Education is the most powerful weapon which you can use to change the world. The power of education extends beyond the development of skills we need for economic success. It can contribute to nation-building and reconciliation."***

# Where AI in Education?

- AI for Learning
  - Personalized learning
  - Diagnosing strengths, weaknesses or gaps in a student's knowledge
  - Providing insights about the progress of a student or class
- AI for Tutoring
  - Adaptive tutors engage students in dialogue, answer questions, and provide feedback
- AI for Testing
  - Generating test questions
  - plagiarism detection
  - **Automatic evaluation of tests**
  - Providing automated feedback

# Education and Assessment

- Evaluating students learning progress is vital process in education




- Popular question type in educational system for both formative and summative assessment

**Objective Question: Choice**

How do you handle missing or corrupted data in a dataset?

- Drop missing rows or columns
- Replace missing values with mean/median/mode
- Assign a unique category to missing values
- All of the above

**Subjective Question: Discussion**

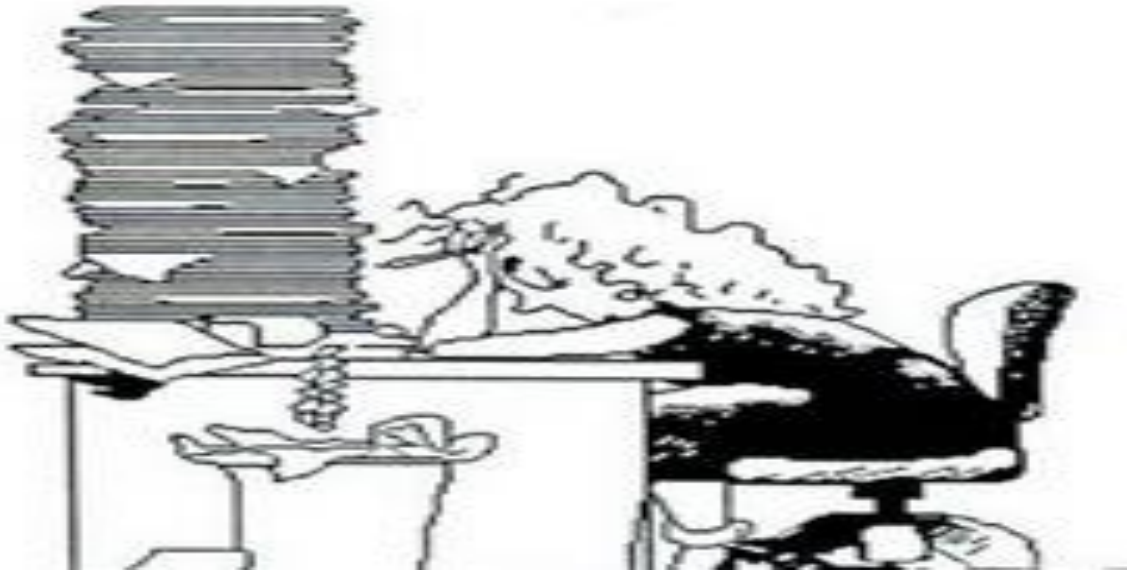
Suppose your EDA is showing the dataset you are dealing with has missing values. Discuss best handling method by considering the type of data point missed and Justify your answer.

**Subjective Question - Essay Type**

Write a persuasive essay to a newspaper reflecting your vies on censorship in libraries. Do you believe that certain materials, such as books, music, movies, magazines, etc., should be removed from the shelves if they are found offensive? Support your position with convincing arguments from your own experience, observations, and/or reading.

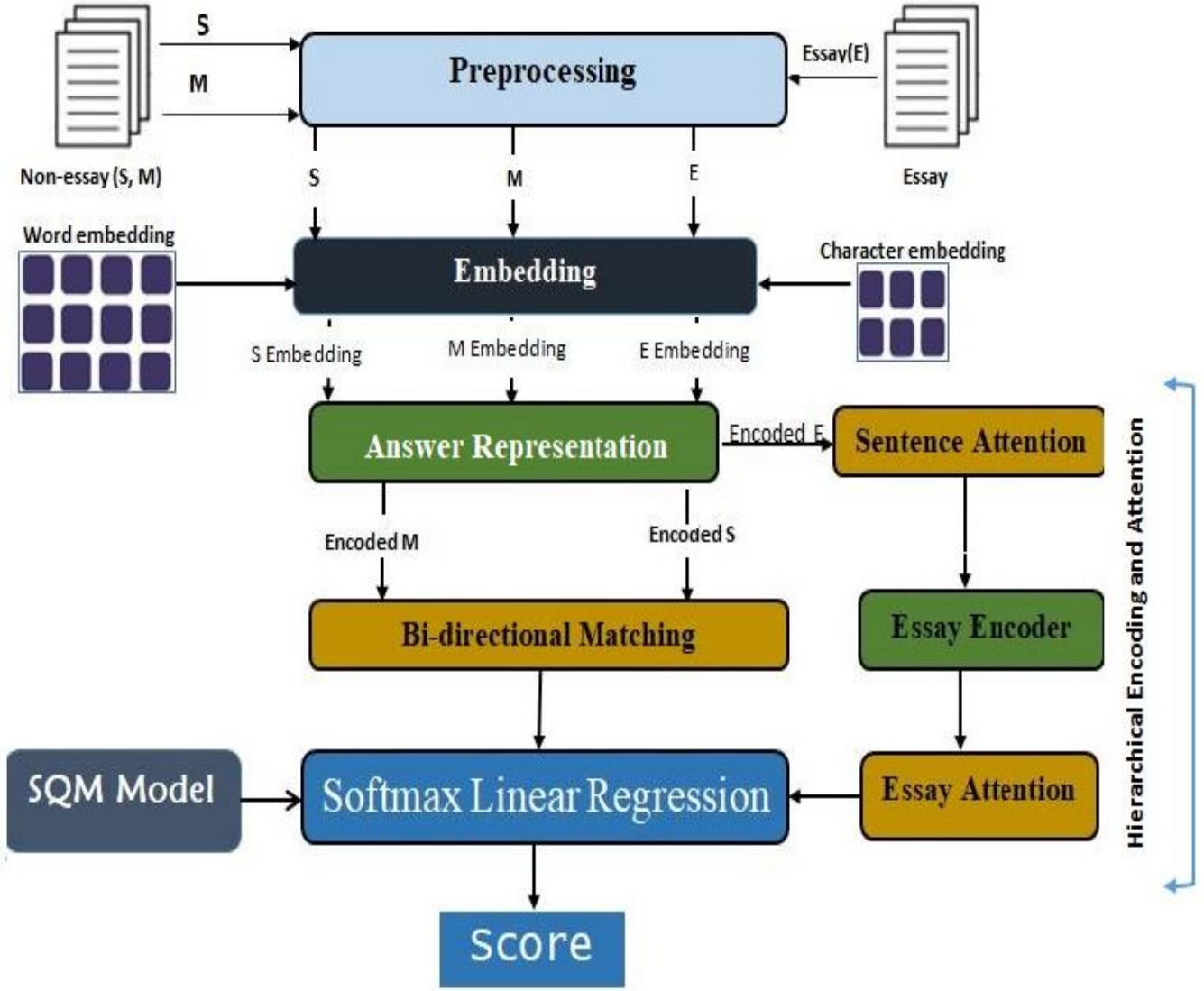
# Education and assessment …

- Manual assessment for subjective question is inherently subjective process
  - Depends on rater personal observation and mood
  - Tedious to evaluate and delayed
  - Might be biased to personal observation



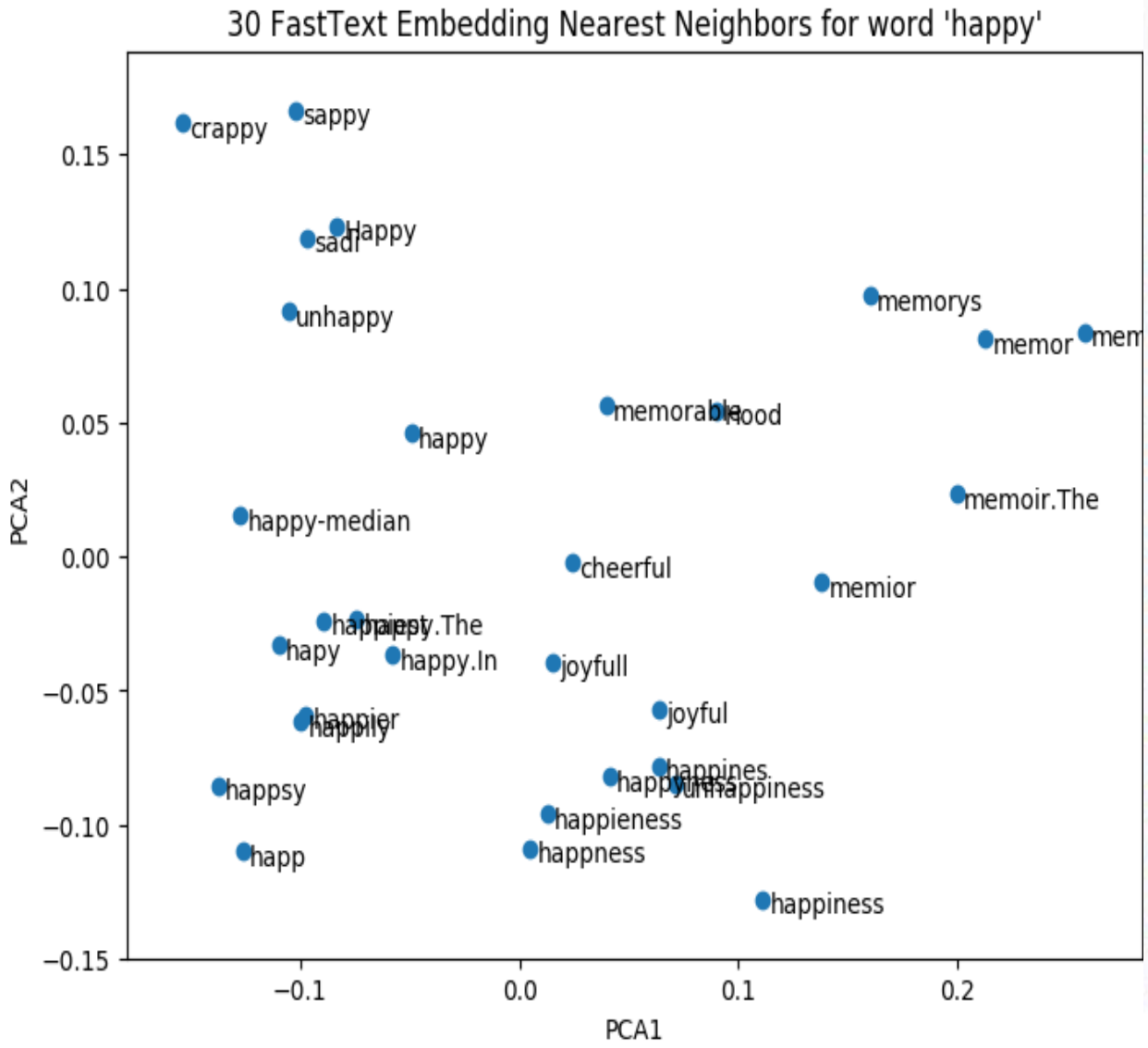
# SQM Architecture

- Preprocessing
- Representing word meaning
- Learning answer level context (Syntax and Semantic Representation)
- Relevancy of texts elements(words, sentences) in an answer, cohesiveness
- Scoring

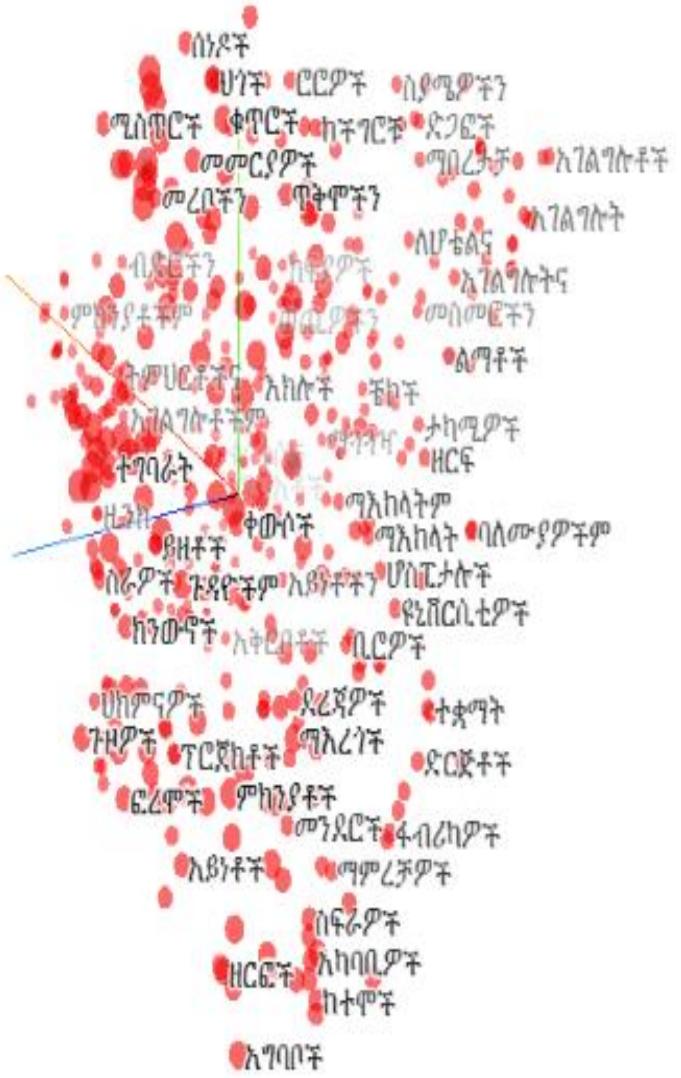


# Word Representation

- **FastText trained on Kaggle essay data**

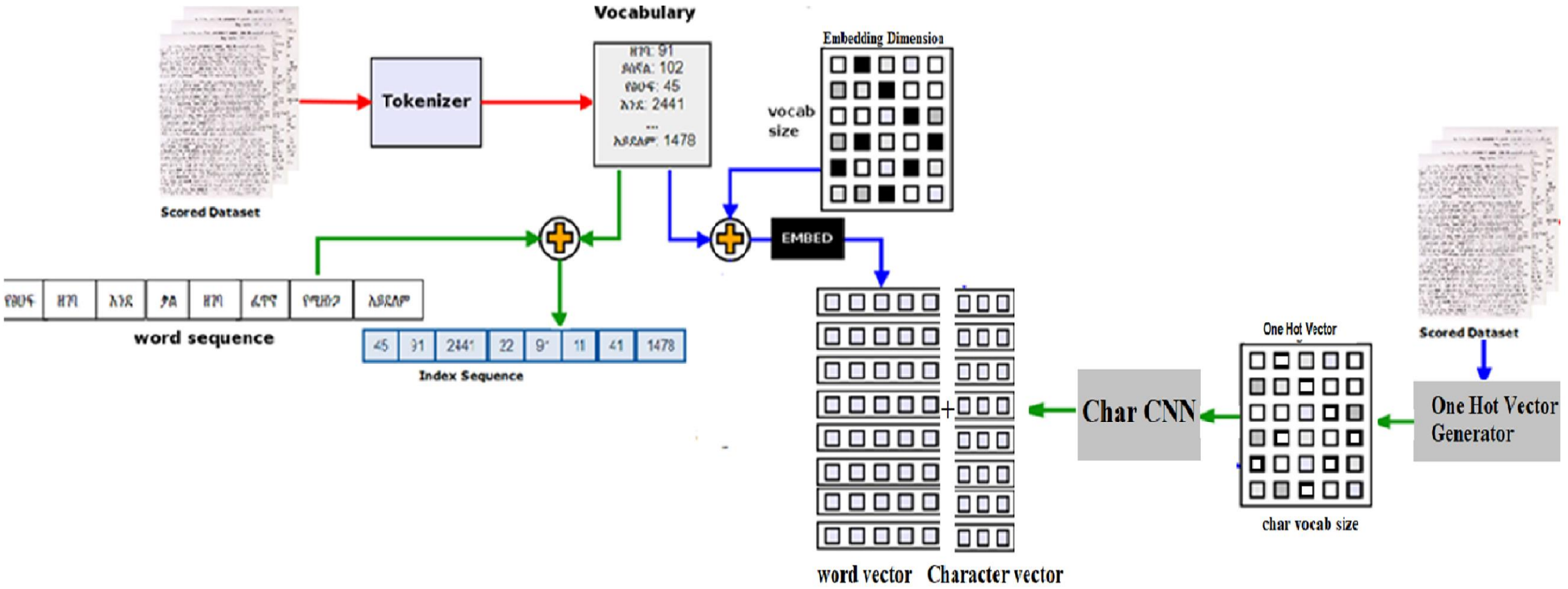



- **FastText trained on Amharic global data (537.5M)**

# Enriching Word Representation with its Sub-word

- Combination of FastText word embedding and convolutional neural network learned character representation used to represent character and word meaning in answer.
  - Supports to learn spell error and representation of OOV words (rare but important)

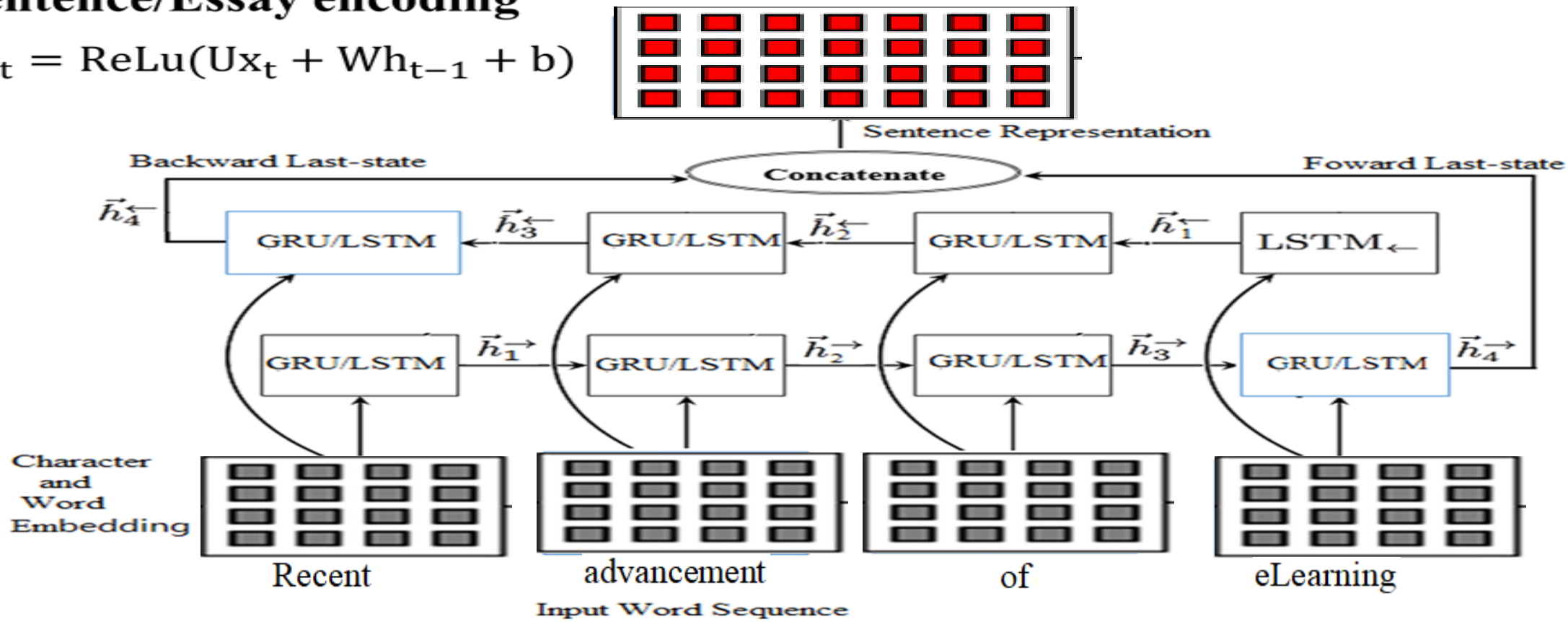


# Answer Representation

- Syntax(grammar) and semantic(meaning) representation of input answer

**Sentence/Essay encoding**

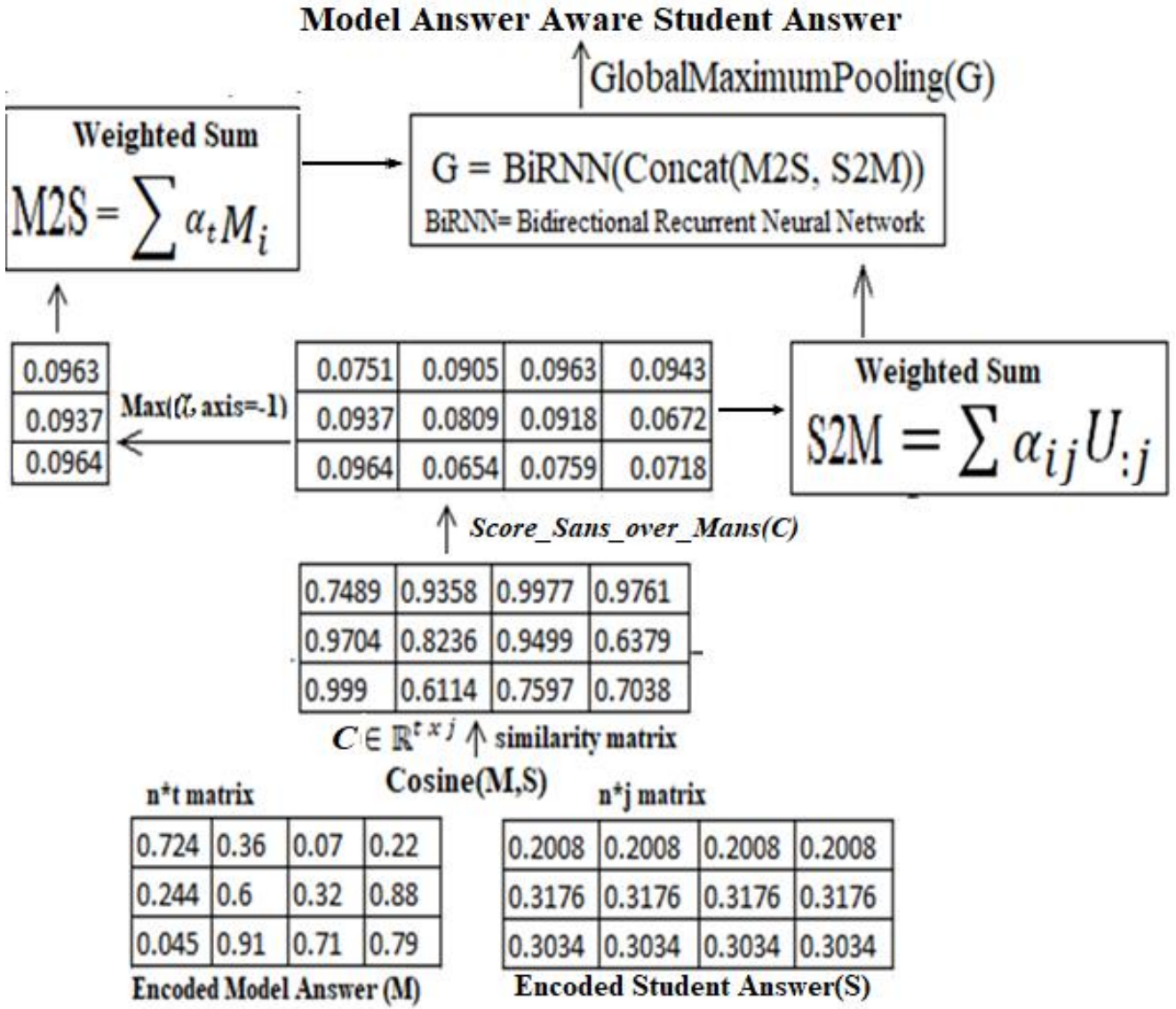
- $h_t = \mathrm{ReLu}(Ux_t + Wh_{t-1} + b)$



# Bidirectional Matching for Non-essay Subjective Questions

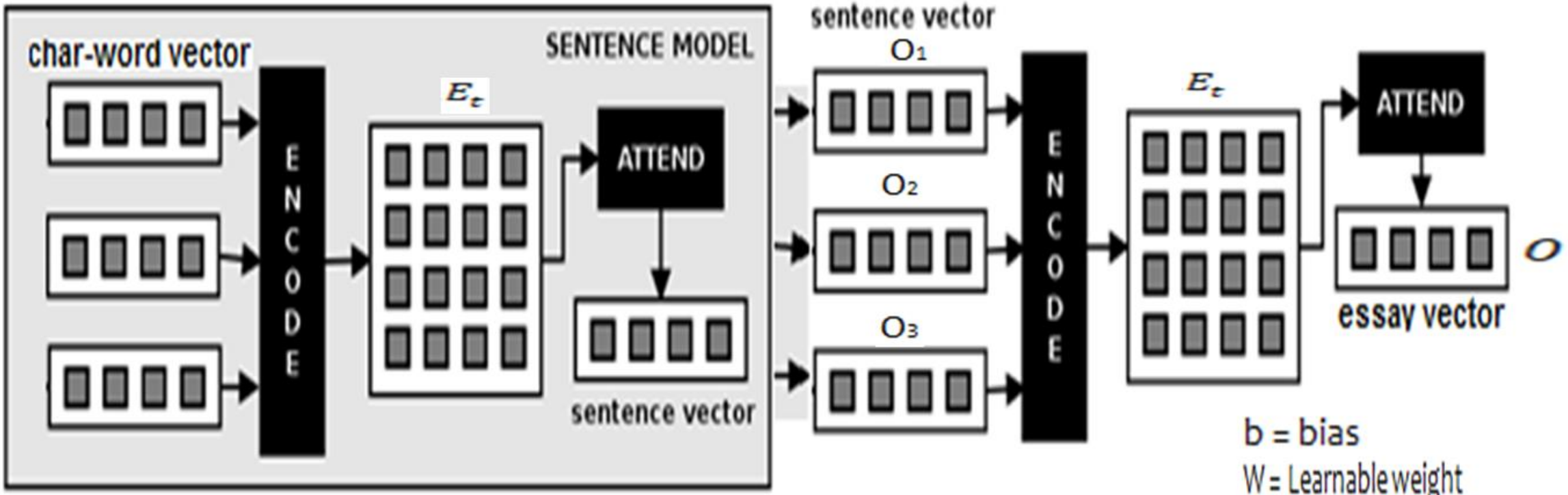
- **Model-answer to Student-answer**: Signifies which student answer words are most relevant to each model answer words

- **Student-answer to Model-answer**: Signifies which model answer words have the closest similarity to the student answer words.



# Hierarchical Attention for Essay Questions

- **Captures intra-essay semantics**: relatedness of words, sentences and paragraphs of an essay
- **Captures essay cohesiveness**: word, sentence and paragraph organization




**Attend** Step description = $R_t = Relu(m_t + WE_t + b)$ $\alpha_t = softmax(R_t)$ $O = \sum \alpha_t E_t$

# Scoring

- Weight of question varies
- NN expects all should lay in fixed range
- Given minimum score $S_{min}$ and maximum score $S_{max}$of the given prompt or question set,
  - calculate model friendly score range that lay between 0 and 1 as:

$$S_i = \frac{S_i - S_{min}}{S_{max} - S_{min}}$$

  , where $S_i$ score for i-$^{th}$ answer in question set.
- For evaluation: reverse score to original range using:

$$S_i = S_i^* (S_{max} - S_{min}) + S_{min}$$

- **Softmax linear regression** used to predict score in defined range

# Dataset

## Dataset for English Essay

- Kaggle dataset is used
- The dataset contains 12,976 essays ranging from 150 to 550 words each, marked by two raters
- state-of-the-art on this dataset has achieved a Cohen's $\kappa = 0.96$ (using quadratic weights)
- Score varying [0-60]

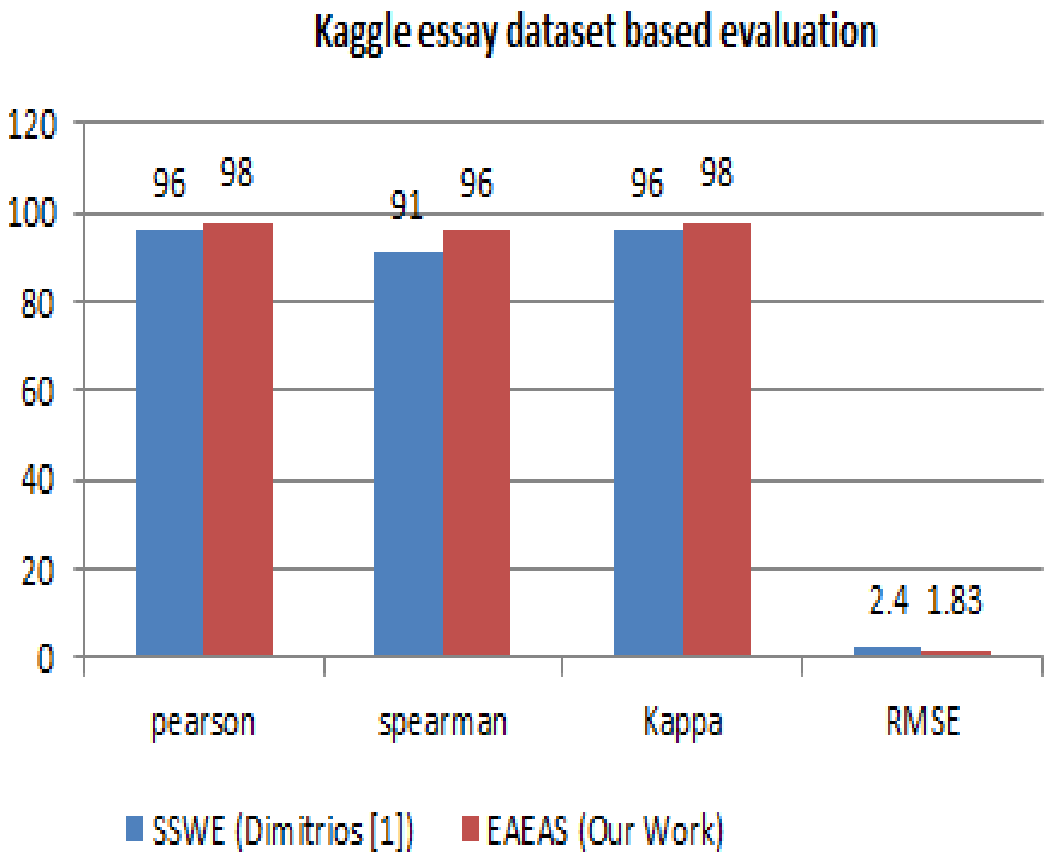
## Dataset for English Short Answer

- Moher released 2442 graded short answer Computer science domain questions
- Reported result:
  - Pearson =0.41, RMSE=1.018
- Weight: 5

## Dataset for Amharic Short Answer

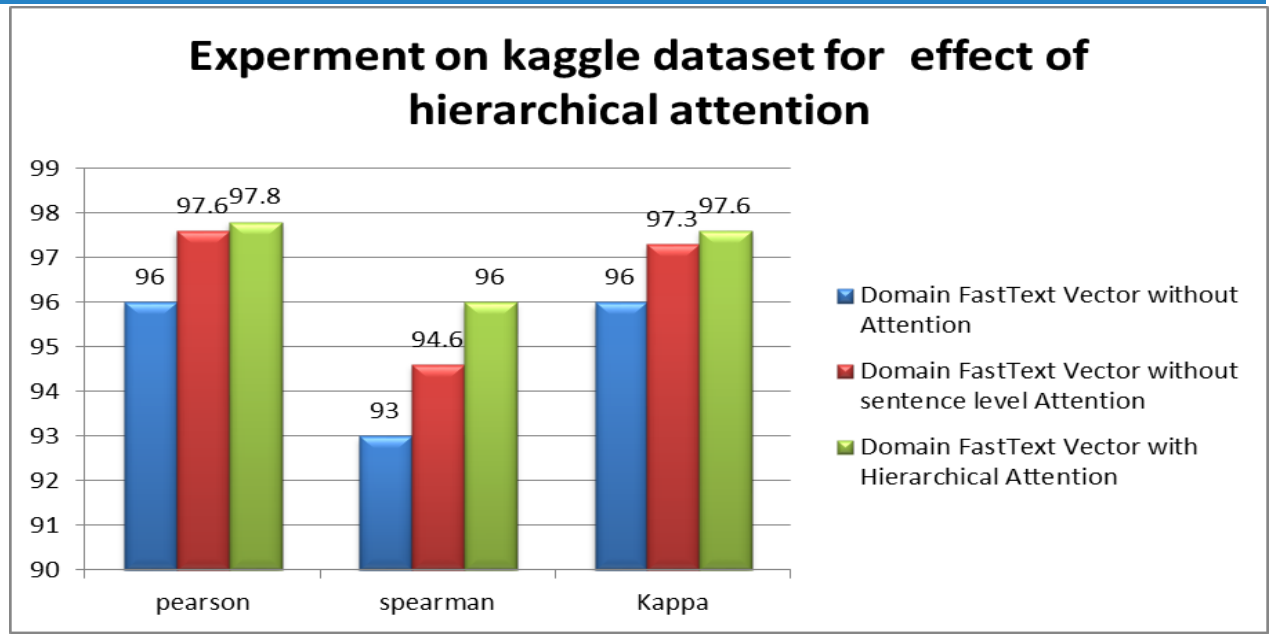
- No standard  dataset is available
- Collected 1112 answers pre-graded
- Rated by two domain raters; Average score is used to assess
- Varying score 0-4; most set weight is [0-3]
- Inter rater correlation between two rates is 87 % Pearson

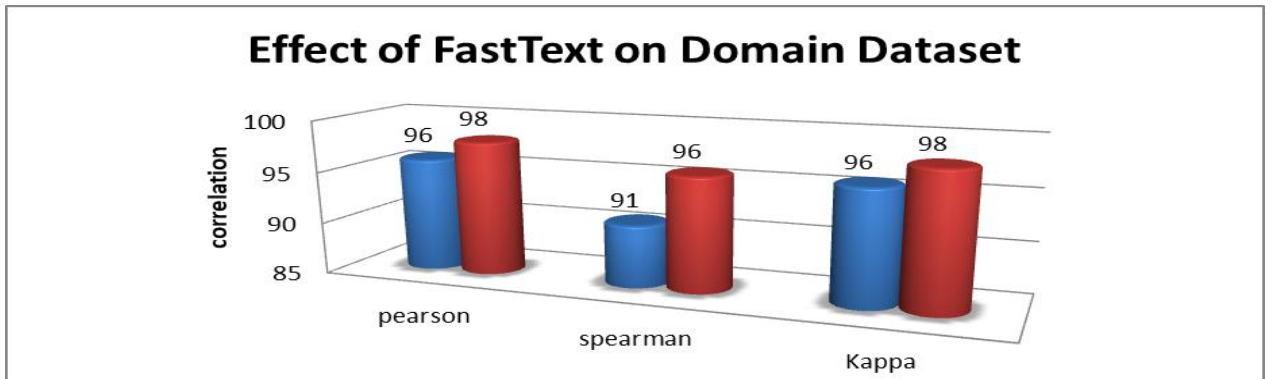
# Result on essay




EAEAS:- Hierarchical encode-attend-encode-attend-score model

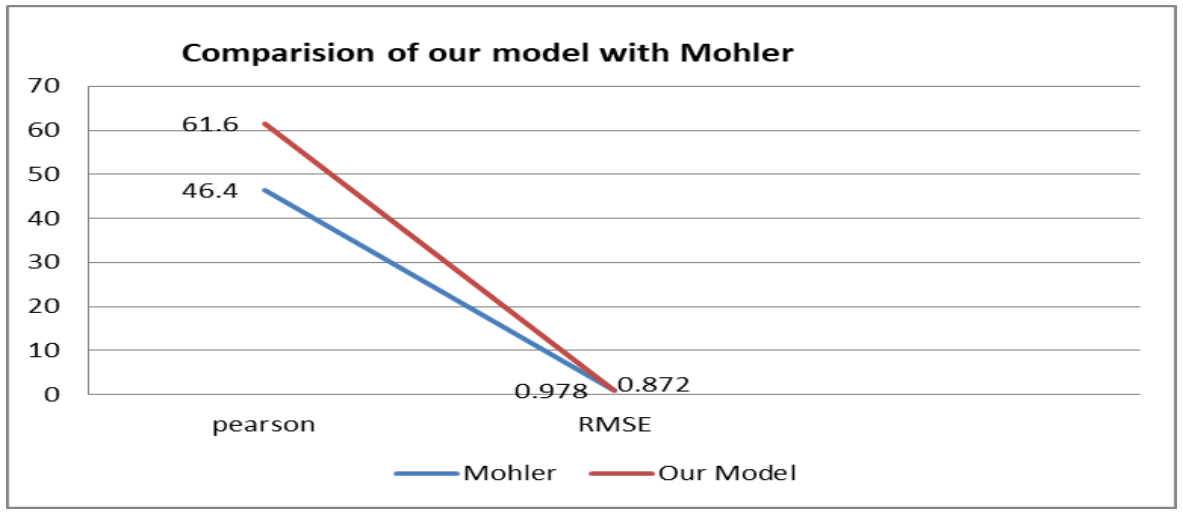
SSWE:- Score Specific Word Embeddings

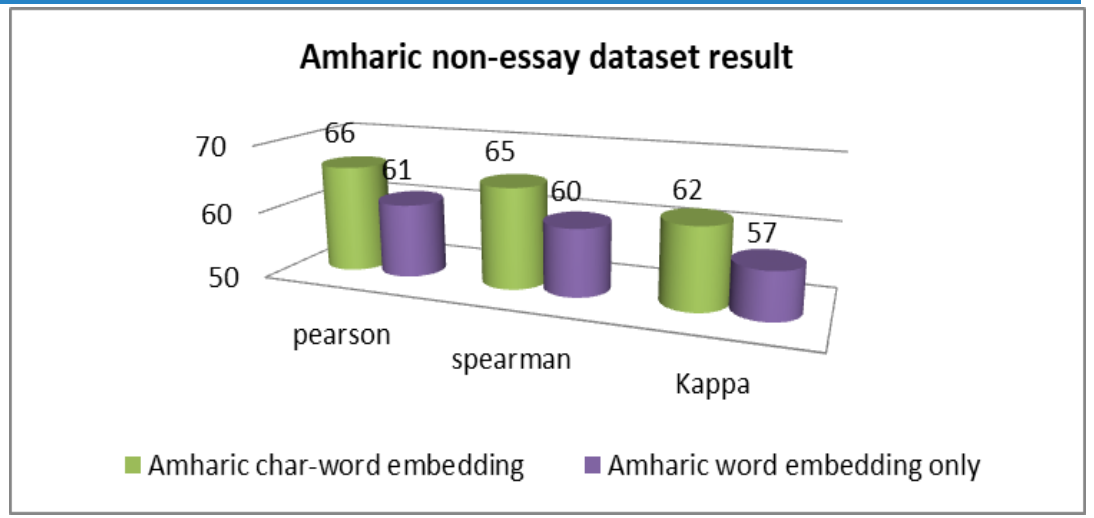


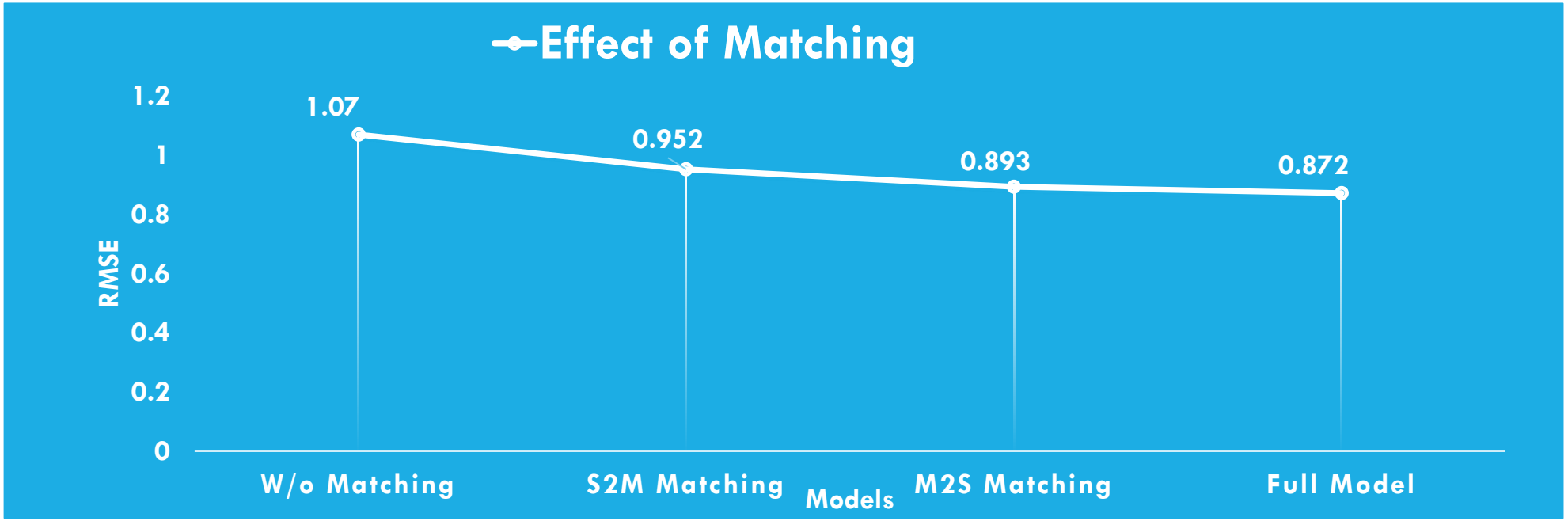




| | pearson | spearman | Kappa |
|---|---|---|---|
| Global Glove vectors with Hierarchical Attention | 96 | 91 | 96 |
| Domain FastText Vector with Hierarchical Attention | 98 | 96 | 98 |

# Result on non-essay



# What is open?

- Questions with graph, table, and figure and mathematical proof type questions
- Feedback with accurate visualization of missed points and justification
- Question item analysis: level of difficulty and coverage to predict reason of failure
- Student answer written recognition: “local languages”
- Educational Chabot: “local languages”
- Learning outcome prediction: personalization
- Student learning behavior analysis
- Clean, large and **inclusive** dataset






# Thank you!

wm2wts@gmail.com, abebaw.eshetu@haramaya.edu.et
https://github.com/Abe2G
https://abe2g.github.io
https://www.linkedin.com/in/abebawu-eshetu
@wm2wts