Evaluation of Answers using Computational and AI Models



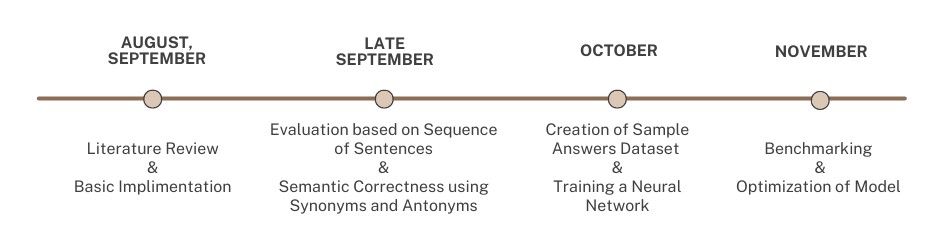
**MOTIVATION**

1. Subjective paper evaluation demands the checker check every word in the descriptive answer for many students
2. This becomes a monotonous and time- consuming process
3. Automatizing of evaluation process makes the process impartial and increases its efficiency



**INTRODUCTION**

The uneven distribution of evaluators and examinees in the society, impartial and favored marking, monotonous and time- consuming evaluation process has created a need for a system that would solve these problems and make the process efficient. Moreover, a subjective answer demands the checker check every word of the answer for scoring actively, and the checker’s mental health, fatigue, and objectivity play a mas- sive role in the overall result. Thus, we are developing a system using the concepts of Natural Language processing that will pro- grammatically evaluate long answers.



**TIMELINE**



**METHODOLOGY**

A crude logic of text-preprocessing and basic answer evaluation was implemented using Python which included:

1. Loading the sample answers from an excel file and loading it into a dataframe.
2. Preprocess the student answer chunk of text: remove stopwords and punctuation and lem- matize the text sentence-wise.
3. For each word in the preprocessed text, we apply the evaluation algorithm which involves checking for

* presence of keywords
* sequence of keywords
* sequence of sentences
* presence of synonyms/antonyms and their effect

1. Allot marks according to a marking scheme decided.
2. Create a dataset that has the set of answers. Train a neural network that can evaluate the answers based on the 2 inputs i.e model answers and student answers.
3. Testing the model and increasing the efficiency.



**RESULTS**

* The crude logic of answer evaluation aided in alloting marks to a text file in the basic way
* The marks alloted were on the basis of presence of keywords and their order
* The total marks for an answer is the sum of the individual marks of each sentence



**CONCLUSION**

1. Generation of keywords, Removal of stop words, Evaluation of answers by comparison of keywords and the rele- vance of the order is done
2. From literature review, we observe that LSTMs are the preferred ways to solve NLP based AI Problems. The quality of the dataset plays a key role in deciding the performance of the model.



**REFERENCES**

**1.**

**2.**

M. F. Bashir, H. Arshad, A. R. Javed, N. Kryvinska, and S. S. Band, “Subjective an- swers evaluation using machine learning and natural language processing,” *IEEE Ac- cess*, vol. 9, pp. 158 972–158 983, 2021.

C. N. Tulu, O. Ozkaya, and U.Orhan, “Auto- matic short answer grading with semspace sense vectors and malstm,” *IEEE Access*, vol. 9, pp.19 270–19 280, 2021.

PROJECT TEAM MEMBERS: NAVYA SIBI, VEDANT PHUSE, DEVDATTA BIDKAR, SNEHAL PATIL

SUPERVISOR NAME: DR. KISHORE BHURCHANDI

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

VISVESVARAYA NATIONAL INSTITUTE OF TECHNOLOGY, NAGPUR-INDIA



**PROBLEM FORMULATION**

We were given a set of problem state- ments like:

1. Fitness Device Biomedical Signal Processing
2. Automatic Subjective Answer Evaluation
3. Image Inpainting and Outpainting

Out of these, we decided to choose the Au- tomatic Subjective Answer Evaluation prob- lem statement, which is essentially Evalua- tion of answers using Computational and AI Models



**LITERATURE SURVEY**

1. Subjective Answers Evaluation Using Machine Learning and NLP [1]
   * In this paper, the answers are evaluated using the solution and provided keywords using various Similarity-based techniques.
   * Two score prediction algorithms are proposed, which produce up to 88% accurate scores. Various similarity and dissimilarity thresholds are studied, and various other measures such as the keyword’s presence and percentage mapping of sen- tences are utilized to overcome the abnormal cases of semantically loose answers.
2. Automatic Short Answer Grading With SemSpace Sense Vectors and MaLSTM [2]
   * This paper makes the use of SemSpace and Manhattan LSTM (MaLSTM), based on multi-layer LSTM Networks for Automatic Short Answer Grading (ASAG).
   * The model was first trained on the open source Mohler dataset
   * A new dataset, CU-NLP was developed for this study, where initial data was con- verted into SemSpace Sense Vectors after which it was trained on a model built on the MaLSTM Network.
   * Two identical LSTMs are trained for Student and Model answers, and each calcu- lates vectorial similarity to find the Manhattan distance which is used to find the similarity value.