Evaluation of Short Answers

Submitted in partial fulfillment of the

requirements for the award of the degree

of

Master of Computer Application

(MCA)

To

Guru Gobind Singh Indraprastha University, Delhi

Guide:

Dr. Sushma Bahuguna

Designtaion: Sr. Ast. Professor

Submitted by:

Hitesh Walia

Rollno: 00311104422

Banarsidas Chandiwala Institute of Information Technology

New Delhi – 110020

Batch (2022-2024)

CERTIFICATE

I, Mr./Ms. Hitesh Walia, Roll No. 00311104422 certify that the Project Report/Dissertation (MCA-269) entitled "Evaluation of Short Answers" is done by me and it is an authentic work carried out by me at Banarsidas Chandiwala Institute of Information Technology, New Delhi. The matter embodied in this project work has not been submitted earlier for the award of any degree or diploma to the best of my knowledge and belief.

Signature of the Student

Date:

Certified that the Project Report/Dissertation (MCA-269) entitled "Evaluation of Short Answers" Done by Mr. Hitesh Walia, Roll No. 00311104422 is completed under my guidance.

Signature of the Guide

Date:

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my project mentor Dr. Sushma Bahuguna for valuable guidance and support throughout the development of the project. I would also like to thank Mr. Alok Mishra for providing valuable inputs and building confidence during project reviews. I would like to express immense appreciation and gratefulness to MCA coordinator of the college, Dr. Sushma Bahuguna for providing the opportunity to embark on the development of this project, and ensuring smooth conduct of various itenaries in the entire process of project development. I would like to thank Mr. Alok Mishra for organizing project reviews and offering valuable help and information regarding development of project. I would like to extend my gratitude to all MCA faculty for being supportive and encouraging, and my class friends for invaluable help, assistance and cooperation in various things.

(Signature)

Name: Hitesh Walia

Roll Number: 00311104422

SYNOPSIS OF THE PROJECT

Evaluation of Short Answers using NLP

1. Introduction

In an era driven by information, educational platforms, and content generation, the need for effective tools to assess and evaluate short-form responses has become increasingly critical. The "Short Answer Evaluation Project" aims to address this need by developing a user-friendly using Natural Language Processing (NLP).

This project seeks to provide educators and learners with a tool that enables the evaluation of short text responses, such as answers to questions, by assigning scores based on their relevance, clarity, and correctness. With the Short Answer Evaluation Project, users can expect improved assessment accuracy.

2. Objective

To Create a user friendly tool that enable users to check their personal accuracy and understanding of a topic or subject matter and provide educators/teachers a convenient way to efficiently evaluate student responses, helping them in grading processes and enhance the quality of education.

3. Name of the Organization

The project is a part of BCIIT third semester Minor Project. Banarsidas Chandiwala Institute of Information Technology (BCIIT) was established in 1999 to run Master of Computer Application (MCA) programme. The Institute is affiliated with Guru Gobind Singh Indraprastha University.

4. Future Scope

Advanced NLP Techniques: Continuously improve the short answer evaluation module by incorporating the latest advancements in natural language processing and machine learning. Explore more advanced algorithms for better understanding and evaluating text responses.

Educational Insights: Develop data analytics and reporting features that provide users with insights into performance trends, helping learners to improve their accuracy and teachers with their teaching strategies.

5. Data Required & Data Collection Method

This project primarily focuses on short answer evaluation, the data collection process involves creating sample questions and their respective answers to build and test the evaluation. The following describes the data requirements and collection methodology:

Data Requirements:

1. ***Sample Questions***: A collection of sample questions of various subjects. These questions serve as input for the short answer evaluation module.
2. ***Reference Answers***: Corresponding reference answers for each sample question. These reference answers are the basis for evaluating the quality of user responses.

Data Collection Methodology:

1. ***Question Generation***: Create a set of sample questions covering different domains and levels of difficulty.
2. ***Answer Creation***: Develop reference answers for each sample question. The reference answers serve as a basis for evaluating user-provided answers.

6. Methodology

The methodology used for Design & Development is as follows:

We will use simple and effective methods to evaluate short descriptive student answers. The similarity between each student's answer with its model answer is evaluated using Natural Language Processing (NLP) techniques. The similarity score is used to calculate the score.

We will pre-process the entered answers by tokenizing, removing noise i.e. unnecessary words that do not have a direct impact on the quality of answer.

User Interface will be made using HTML, CSS and Bootstrap and Flask will be used for the integration and functioning of the application.

7. Hardware Requirements

Type of Hardware	Hardware requirements
Hardware	Intel core ® 2.70 GHZ processor .64-bit system
Installed Memory(RAM)	4 GB or higher

8. Software Requirements

Types of Software	Software Requirements
Operating System	Windows 7 or higher
Web browser	Google Chrome or FireFox
Code Editor	Visual Studio Code
Python Libraries	Flask, Scikit-Learn, NLTK

EVALUATION OF SHORT ANSWERS

Chapter-1: Introduction/Problem Definition

1.1 Name of the Organization:

The project is a part of BCIIT third semester Minor Project. Banarsidas Chandiwala Institute of Information Technology (BCIIT) was established in 1999 to run Master of Computer Application (MCA) programme. The Institute is affiliated with Guru Gobind Singh Indraprastha University.

1.2 General Description of the System under Study:

Education is fundamental for human progress. A student is evaluated by the mark he/she scores. The evaluation of student's work is necessary for analyzing his/her performance and is also a central aspect of the teaching profession that can affect students in significant ways.
The Short Answer Evaluation system is designed to address the evaluation needs of both students and teachers, providing distinct benefits from each perspective.

Students can use the system to assess the accuracy and quality of their short text responses to questions or prompts. This enables learners to identify areas for improvement.

Though teachers use multiple criteria for assessing student work, it is not known if emotions are a factor in their grading decisions. Also, there are several mistakes that occur on the department's side like totaling error, marking mistakes. So, we are developing software to automate the evaluation of answers using Natural Language Processing, significantly reducing the time and effort required for manual grading.

1.3 Establish the need of New System:

The score given for short answers may vary from instructor to instructor. There are many short answer grading and essay grading systems existing that are either automated or semi-automated. Automated grading systems reduce human effort, but no popular tool still exists. In this project, we are focusing on short-answer grading systems. We use simple and effective methods to evaluate short descriptive student answers. The similarity between each student's answer with its model answer is evaluated using word embedding algorithms. The similarity score is used to calculate the score.

1.4 Objective:

To Create a user friendly tool that enable users to check their personal accuracy and understanding of a topic or subject matter and provide educators/teachers a convenient way to efficiently evaluate student responses, helping them in grading processes and enhance the quality of education.

1.5 Methodology:

The methodology used for Analysis, Design & Development is as follows:

- The coding part of is done in Visual Studio Code (VS Code), which is a free and open-source code editor developed by Microsoft. It is immensely popular among developers for its flexibility, speed, and wide range of extensions and features.
  It includes pre-processing the user answers to extract relevant information which involves steps such as Tokenization, Removing Stopwords, Lemmatization. Then we will evaluate the answers by using vectorization and calculating the cosine similarity score between the reference answer and the user's answer.

- The User Interface for entering the required details (answers) and displaying the evaluation result is made using Bootstrap, HTML and CSS which provides an easy and attractive interface to the user.

- The local machine deployment is done in Flask which is a web application framework written in python, it helps end users interact with your python code directly from their web browser without needing any libraries, code files, etc.

Libraries such as NLTK (Natural Language Processing Toolkit), Scikit-Learn, Flask would be required to accomplish these tasks.

1.6 Data Required & Data Collection Method

This project primarily focuses on short answer evaluation, the data collection process involves creating sample questions and their respective answers to build and test the evaluation. The following describes the data requirements and collection methodology:

Data Requirements:

3. ***Sample Questions***: A collection of sample questions of various subjects. These questions serve as input for the short answer evaluation module.
4. ***Reference Answers***: Corresponding reference answers for each sample question. These reference answers are the basis for evaluating the quality of user responses.

Data Collection Methodology:

3. ***Question Generation***: Create a set of sample questions covering different domains and levels of difficulty.
4. ***Answer Creation***: Develop reference answers for each sample question. The reference answers serve as a basis for evaluating user-provided answers.

Chapter-2: System Requirement Analysis

2.1 Identification of Processes:

In this section, we identify and describe the processes that constitute the core functionalities of the Short Answer Evaluation System. These processes form the backbone of the system and include:

- Setting questions and their reference answers.
- User input (answers) to the questions.
- Calculation of cosine similarity score between the reference answer and user's answer.
- Displaying the evaluated score to the user.

2.2 Input and Output Identification:

Here we are dealing with textual data in the form of questions and answers and Natural Language Processing is most suitable for processing the textual data. So, the inputs are in the form of text and output is calculated as a numerical score based on the quality of user's answer.

2.3 Procedures and Rules:

As mentioned earlier, we will be dealing with textual data so it will involve a series of steps from taking user input to displaying the evaluated score to the user.

(a)Data Pre-Processing: It involves of cleaning of data to extract relevant information. It will include the following steps:

- ***Tokenization***: Tokenization is the process of splitting a text into smaller units, which are often words known as **tokens**. In Natural Language Processing (NLP), tokenization is a crucial preprocessing step before analyzing or processing textual data. Tokens are the building blocks of text analysis, and they help facilitate various NLP tasks, such as text classification, sentiment analysis, machine translation, and more.

- ***Stopwords Removal****:* Removing stopwords helps reduce the "noise" in text data, it involves the elimination of commonly used words (stopwords) from text before further analysis or processing. Stopwords are words that are generally considered to be of little value or importance in text analysis because they are very frequent in most texts and do not carry significant meaning on their own.

  Examples of common stopwords in English include words like "the," "a," "an," "in," "and," "of," "to," "for," "is," "with," and so on.

  However, stopwords removal is not always appropriate or necessary for all NLP tasks. In some cases, stopwords might carry essential information.

- ***Lemmatization****:* It is a NLP technique used in text analysis to reduce words to their base or root forms. It reduces words to their canonical or dictionary form, known as the "lemma." It considers the context and part of speech of the word to ensure that the resulting lemma is a valid word and maintains its meaning.

  For example, "jumping" would be lemmatized to "jump."

(b) Evaluation Part: This will include converting textual data into vectors and calculating the cosine similarity between the answer and the reference answer

- ***Vectorization*****:** It refers to the process of converting textual data, such as words or documents, into numerical vectors or arrays of numbers. This numerical representation is necessary because machine learning algorithms and models typically work with numeric data, and text data is inherently non-numeric. Vectorization allows you to represent text in a format that can be processed by various machine learning and statistical models.

  There are several techniques for text vectorization in NLP, including One Hot Encoding, Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF).

- ***Cosine Similarity Score:*** Cosine similarity is a measure of similarity between two non-zero vectors that measures the cosine of the angle between them. In the context of natural language processing (NLP) and information retrieval, cosine similarity is often used to determine the similarity between two text documents or vectors representing text data. It is particularly useful for tasks like document similarity, information retrieval, and recommendation systems.

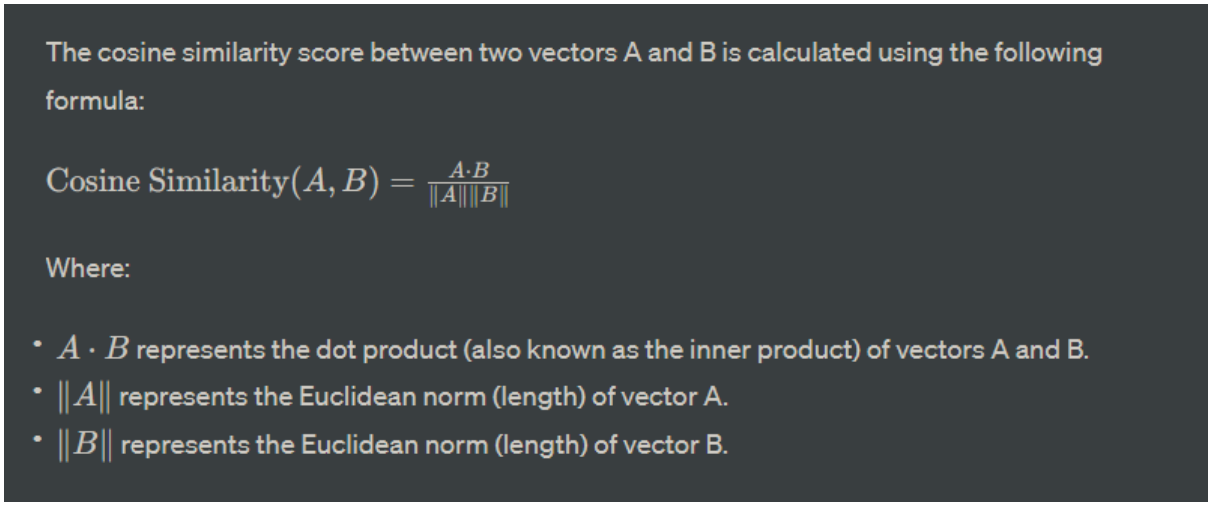

Figure 2.1: Cosine Similarity formula

The resulting cosine similarity score is a value between -1 and 1, where a higher score indicates greater similarity between the two vectors. A score of 1 indicates that the vectors are identical, a score of 0 means they are orthogonal (no similarity), and a score of -1 indicates dissimilarity in the opposite direction.

2.4 System Requirement Tools:

- **Hardware Requirements**

Table 2.1: Hardware Requirements

Type of hardware	Hardware requirements
Hardware	Intel core ® 2.70 GHZ processor .64-bit system
Installed Memory [RAM]	4 GB or higher

- **Software Requirements**

Table 2.2: Software Requirements

Type of software	Software requirements
Operating System	Windows 7 or higher
Web browser	Google Chrome or FireFox
Code Editor	Visual Studio Code
Model Deployment	Flask
Python Libraries	NLTK, Scikit-learn, Flask