

**A Synopsis
On**

Virtual Self Assessment Model

Minor Project for Partial Fulfillment of the Requirements

for the Degree of

Bachelor In Technology

in

Computer Science and Engineering

By

Puranjay Makhija(05415002716)

Pranjal Rastogi(05215002716)

Hemant Garg(03215002716)

Under Supervision of

Ms. Shalu, Assistant Professor CSE



COMPUTER SCIENCE AND ENGINEERING
MAHARAJA SURAJMAL INSTITUTE OF TECHNOLOGY

C – 4, JANAKPURI, NEW DELHI - 110058

August, 2019

Abstract

The project will tackle the issue of lack of a practice platform for students to prepare for video interviews and essay writing for interview/exam preparation. Our platform will act as a self-assessment platform for students to practice and judge themselves with the help of built-in machine learning models and functionalities to simulate real environment and scoring.

The project will aid in student practice through functionalities of a live interview and essay writing for exams.

Our platform will help students judge the accuracy and proficiency of their essays, as well as check for plagiarism simultaneously so as to avoid any issues in the real exam. Automated real-time scoring will be implemented for students, which will help them understand the key points helpful for increasing their scores over a period of time.

Live Interview functionality will help students gain confidence before actual interviews, and will help to prepare answers. A confidence score will be calculated to help students improve. Also, students' answers will be converted to text and will be sent to them, so that they can assess their answers better and improve further.

Introduction To Project:

Our project includes two major functionalities:

- 1) Essay Writing Analysis
- 2) Live Interview audio analysis

1) Essay Writing Analysis:

Through text analysis, we will judge the accuracy and proficiency of the essay written by the user and gave them an automated score. This analysis will be done through a deep learning model created by us. The report generated will also be sent to the user with accuracy predicted to their email ids.

Another feature in text analysis is to include plagiarism detection which will prevent a user from copying an essay from other online resources. The technique for plagiarism detection will be done through the Machine Learning model which will return the percentage of similar text.

2) Live Interview Audio Analysis:

In the audio analysis, our project mainly focuses on voice to text and voice sentiment analysis. In voice sentiment analysis our model will predict different emotions in the voice of the user and will generate a confidence score based on these factors.

In the audio analysis, we are also be adding the feature of converting voice to text so as to help users, self asses their answers for future use. The answers will be mailed to their respective email id in the form of text. After performing voice to text conversion, our algorithm will also detect the number of repetitive words that will be highlighted, so as to help users asses and avoid the same in a real interview.

Literature Survey:

We reviewed several papers and blogs to improve our understanding of models used for various functionalities.

Plagiarism detection- We reviewed several techniques like fingerprinting, bag of words model and tools and API available. We are moving forward with affinity-based, frequent analysis model, using string sub matching and word frequency model, to detect plagiarism, even if, some paraphrasing is performed[2].

Automated Scoring- A quadratic weighted kappa model, to measure agreement between 2 different rates, namely LSTM and relu function over 5 fold epoch strategy will be used to improve accuracy[1][3].

Speech to text- A recurrent neural network-based approach, to improve accuracy over time will be used. In this Recurrent Neural Network(RNN) along with Connectionist Temporal Classification(CTC). The role of CTC is to make a sequence of repetitive characters without space into one. In RNN output is 10 times smaller than the input[4].

Audio Analysis- Support Vector Machines(SVM) are used for performing sentiment analysis on audio. It is a non-probabilistic model that uses a representation of text examples as points in multi-dimensional space.

Objectives:

1. To help students prepare for interviews and exams, through real environment simulation and scoring. Students can assess their progress, through various automated tools and functionalities and can improve.
2. Automated essay scoring and plagiarism detection to help students prepare their answers for exam preparation/application for higher studies.
3. Time constraints on answering and confidence analysis through voice, so as to help students gain prepare confidence for real interviews.
4. Text conversion of their answers, to help students assess their progress.

Our platform will act as a practice platform for students to assess their progress through time-constrained, real-world simulated environment and automated scoring facility to help students prepare better for their interviews/exams. Our target users will be students preparing for such activities.

Research Methodology

The project will comprise of 2 major options, to either work on essay writing or live interview. The editor will be provided for essay writing and a live recorded session facility for interview practice, with a random question from our questions dataset and a timer, which will be coordinated on the server side. The recorded answer/text will be then sent to the database, hashed by user email id, and processing will be done on the server side, based on our models. The final output will be then sent to users, either real-time or on their email ids.

As mentioned above our project include two types of analysis:

- 1) Text Analysis
- 2) Speech Analysis

Text Analysis:

In text analysis, we will perform essay writing analysis and implement the automated scoring system and plagiarism detection.

The model used for automated scoring system consists of 2 Long Short Term Memory(LSTM) layers with a Dense output layer. The final layer uses the Relu activation function.

The model used for analysing essay is Frequent-itemset analysis, using string matching.

Speech Analysis

Speech analysis includes converting speech to text and doing voice sentiment analysis.

The Model used for converting speech to text is Recurrent Neural Network(RNN)

The model used for voice sentiment analysis is Support Vector Machines(SVM)

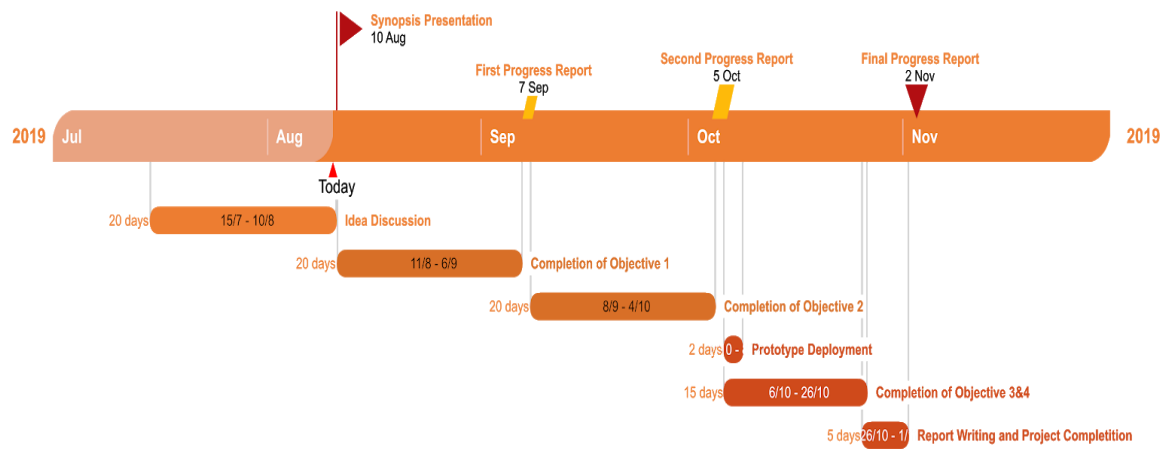
Tech Stack:

The following technologies will be used for making our project work.

1. ReactJS
2. NodeJS
3. Python
4. Django
5. MongoDB
6. Git
7. 'Machine Learning' and 'Deep Learning' libraries

Timeline:

The timeline for project completion is as follows.



References

[1] Dimitrios Alikaniotis and Helen Yannakoudakis, "Automatic Text Scoring Using Neural Networks ", arXiv:1606.04289v2 [cs.CL] 16 Jun 2016

[2] Hussain A Chowdhury and Dhruba K Bhattacharyya, "Plagiarism: Taxonomy, Tools and Detection Techniques", 19th National Convention on Knowledge, Library and Information Networking (NACLIN 2016), 1801.06323

[3] Kaveh Taghipour and Hwee Tou Ng, "A Neural Approach to Automated Essay Scoring" ,2016 Conference on Empirical Methods in Natural Language Processing, pages 1882–1891,

[4]Suvro Banerjee, "An Introduction To Recurrent Neural Networks".
<https://medium.com/explore-artificial-intelligence/an-introduction-to-recurrent-neural-networks-72c97bf0912>