EDUEASY – SMART LEARNING ASSISTANT SYSTEM

Jayasuriya J.A.T.P.

(IT17112192)

Dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Hons) in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

September 2020

DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also I hereby grant to Sri Lanka Institute of Information Technology, the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:	Date:
The supervisor/s should certify the dissertati	on with the following declaration.
The above candidate has carried out resear under my supervision.	ch for the bachelor's degree dissertation
Signature:	Date:

ABSTRACT

Usage of smart learning concepts has been increased rapidly all over the world recently as better teaching and learning methods. Most of the educational institutes such as universities are experimenting those concepts with their students. Smart learning concepts are especially useful for students to learn better in large classes.

In large classes, lecture method is the most popular method of teaching. In lecture method, the lecturer presents the content mostly using lecture slides and the students make their own notes based on the content presented. However, some students may find difficulties of the above method due to various issues such as speed in delivery.

The purpose of this research is to assist students in large classes in the following content. The research proposes a solution with four components namely note taker, slide matcher, reference finder and question presenter which are helpful for the students to obtain a summarized version of the lecture note, easily navigate to the content and find resources, and revise content using questions.

ACKNOWLEDGEMENT

Working on this project on "EduEasy - Smart Learning Assistant system" was a source of immense knowledge for us. We would like to express my thanks of gratitude to all those who provided me the possibility to complete this project. A special gratitude we give to Dr. Janaka Wijeykoon, the Lecture-in-charge of Comprehensive Design and Analysis Project, Dr. Anuradha Karunasena, our project supervisor & Ms. Pradeepa Bandara, our co-supervisor, whose contribution in stimulating suggestions and encouragement helped us to complete the research project.

The guidance and support received from our families, friends and fellow group members was more helpful to finish the project successfully. Furthermore, we would like to acknowledge all the others who offered their fully support on completion of this research project.

TABLE OF CONTENTS

DECLARATION	i
ABSTRACT	ii
ACKNOWLEDGEMENT	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	vi
LIST OF ABBREVIATIONS	vii
1. INTRODUCTION	1
1.1. Background Literature	1
1.2. Research Gap	6
2. RESEARCH PROBLEM	7
3. RESEARCH OBJECTIVES	8
3.1. Main Objective	8
3.2. Specific Objectives	8
4. METHODOLOGY	9
4.1. Methodology	9
4.1.1. System Architecture	12
4.1.2. User Interface Implementation	13
4.1.3. Project Management	15
4.1.4. Development	18
4.1.5. Software System Attributes	19
4.1.6. Operations	21
4.1.7. User Characteristics	21
4.2. Commercialization Aspects of the Product	22
5. TESTING & IMPLEMENTATION RESULTS & DISCUSSION	23
5.1. Testing	23

5.2.	Results	27
5.3.	Research Findings	29
5.4.	Discussion	29
6. CC	ONCLUSION	30
7. RE	EFERENCES	31
8 AP	PPENDICES	32

LIST OF FIGURES

Figure 1: NLP Processes of the Note Taker component	11
Figure 2: System Diagram of the Note Taker	12
Figure 3: Speech-to-Text Converter UI	13
Figure 4: Text Summarizer UI	13
Figure 5: Summary Display UI	14
Figure 6: Project Management	15
Figure 7: Agile Development	17
Figure 8: Python Programming Language	18
Figure 9: Anaconda Spyder IDE	18
Figure 10: Flask Framework	19
Figure 11: Results of the survey	27
Figure 12: Gantt Chart	32

LIST OF ABBREVIATIONS

Abbreviation Description

NLP Natural Language Processing

NLTK Natural Language Toolkit

UI User Interface

API Application Programming Interface

TF Term Frequency

SDK Software Development Kit

SDLC Software Development Life Cycle

WER Word Error Rate

HITS Hyperlink-Induced Topic Search

1. INTRODUCTION

1.1. Background Literature

The smart learning concepts have used rapidly all over the world for better teaching and learning methods [1]. Smart learning reflects how advanced technologies are enabling learners to digest knowledge and skills more efficiently and conveniently. Every day, researchers are seeking out for new technologies and new methods to improve smart learning concepts. So the researchers inclined towards smart learning concepts are growing more than ever before. Although the ongoing global movement towards smart learning makes and important revolution in modern learning, many people still find these concepts vague [2].

Society is normally resistant to change. Since smart learning is in developing stages, it faces several barriers such as rigid nature of traditional education, insufficient time to learn e-learning tools and implement them into teaching practice, lack of recognition, teachers' low self-efficiency in educational technology knowledge, shortage of knowledge and motivation among students [3] etc. However, smart learning can offer many benefits to the academic community. These benefits include increased productivity [4], interactive and enhanced learning experience, simplified teaching, cost reduction etc. The academic community and public at large, should embrace smart learning as solution to many problems.

There are several issues when it comes to a traditional lecture room environment. Sometimes, writing a note in a lecture room environment is difficult because of several reasons [5]. Sometimes the lecturers are speaking too fast. Sometimes students are focused on trying to understand the lecture so they miss to write the note. Since note taking is hard work, transcribing lecturer's voice using a specifically designed program could be explored as a solution.

The component designed for note taking has two main parts. In the first part the component, voice of the lecturer is transcribed to text. There are some previous

researches done on this voice transcription part [6]-[7]. In those researches, they use different methods to transcribe audio into text.

The research "Automatic Transcription of Spontaneous Lecture Speech" is a very extensive project [6]. The project was conducted over five years in pursuit of fulfilling three major targets which were building a large scale spontaneous speech corpus, acoustic and linguistic modeling for spontaneous speech recognition and constructing a prototype of a spontaneous speech summarization system. They have trained initial baseline models using a large corpus of lecture presentations & talks and confirmed significant difference of real lectures & written notes. They have proposed sequential decoding and speaking-rate dependent decoding strategies because in spontaneous lecture speech, the speaking rate is generally faster and changes a lot. The sequential decoder simultaneously performs automatic segmentation and decoding of input utterances. Then they have applied the most adequate acoustic analysis, phone models and decoding parameters according to the current speaking rate. Those strategies achieved improvement on automatic transcription of real lecture speech.

"Web-Based Language Modelling for Automatic Lecture Transcription" is a study [7] that propose a Language Model (LM) for lecture transcription which eliminates the need for multiple models, yet achieves similar or better Word Error Rate (WER) reduction. Internet broadcasting is becoming an increasingly popular method of delivering lectures and academic presentations. However, without transcripts, users of online lectures are faced with great difficulties in performing tasks. Automatic Speech Recognition (ASR) is a cost-effective method to deliver transcriptions, but its accuracy for lectures is not satisfactory. Most lecture recognition systems achieve Word Error Rates (WERs) of about 40-45% [7], quite far from the minimum WER of 25% for a transcript to be useful and accepted by users.

The next part of the component is the text summarization. Automatic text summarization is a complex task which contains many sub-tasks in it. Every subtask has an ability to get good quality summaries. There are also researches done on automatic text summarization, text summarization machine learning algorithms etc. [8]-[10]. They use methods like machine learning algorithms and clustering & extraction to perform text summarizations.

"Graph-Based Algorithms for Text Summarization" is a research [8] that presents innovative unsupervised methods for automatic sentence extraction using graph-based ranking algorithms and shortest path algorithm. Graph-based ranking algorithms are essentially a way of deciding the importance of a vertex within a graph, based on information drawn from the graph structure. In this study, they have used two graphbased ranking algorithms which were previously found to be successful on a range of ranking problems. Those algorithms are Hyperlink-Induced Topic Search (HITS) and PageRank algorithms. They can be adapted to undirected or weighted graphs, which are particularly useful in the context of text-based ranking applications. HITS algorithms is a link analysis algorithm that rates web pages. It determines two values for a page, its authority (value of the content of a page) and its hub value (value of its links to other pages). PageRank algorithm is also a link analysis algorithm used by the Google internet search engine that assigns a numerical weighting to each element of a hyperlinked set of documents with the purpose of measuring its relative importance within the set. According to this research, Shortest-path algorithm is better because it generates smooth summaries as compared to ranking algorithms.

In comparison, TextRank works well because it does not only rely on the local context of a text unit, but rather it takes into account information recursively drawn from an entire graph. Through the graphs it builds on texts, TextRank identifies connections between various entities in a text and implements the concept of recommendation. An important aspect of TextRank is that it does not require deep linguistic knowledge, domain or language specific annotated corpora, which makes it highly portable to other domains, genres or languages. On the other hand, the Shortest-path algorithm is easy to implement and relatively language independent. Researchers states that the summaries they generated using Shortest-path algorithm are often somewhat smooth to read. When it comes to including the important facts from the original text, the weighting of sentences using traditional extraction weighting methods seems to be the most important part and this Shortest-path algorithm has performed it well in tests. After an extensive comparison, the researchers had come into conclusion that the Shortest-path algorithms is better because it generates smooth summaries as compares to ranking algorithms.

In the research "A Frequent Term and Semantic Similarity based Single Document Text Summarization Algorithm" [9], a frequent term based text summarization algorithm is designed and implemented in java. This algorithm was implemented using open source technologies like java, DISCO, Porters stemmer etc. and verified over standard text mining corpus. According to the research, this designed algorithms works in three steps. In the first step, the document required to be summarized is processed by eliminating stop words and by applying stemmers. In the second step, term-frequent data is calculated from the document and frequent terms are selected and for these selected words, the semantic equivalent terms are also generated. Finally in the third step, all the sentences in the document, which are containing the frequent and semantic equivalent terms are filtered for summarization.

According to the research paper [9], semantic similarity is a concept whereby a set of documents or terms within term lists are assigned a metric based on the likeness of their meaning/semantic content. Various semantic similarity techniques are available which can be used for measuring sematic similarity between text documents. Semantic similarity methods are classified into four main categories which are Edge Counting Methods, Information Content Methods, Feature Based Methods and Hybrid Methods. In this research, semantic similarity of frequent terms is also used to preserve the meaning of original text document in the summarized document.

In this research, the system is divided into three parts which are an input text document, a summarizer algorithm and a summarized text document as output. The summarizer algorithm is further divided into three parts which are the text processing module, frequent terms generation module and the semantically similar terms and sentence filtering module for summarization. The algorithm takes two input parameters, the input text document and number of frequent terms. As the output, it generates a summarized text document along with the two measures compression ratio, which means how much shorter the summary is than the original and retention ratio, which means how much of the central information is retained.

The research "Extractive Text Summarization Using Sentence Ranking" [10] demonstrates a novel statistical method to perform an extractive text summarization on a single document. In this attempt, sentences are ranked by assigning weights and they are ranked based on their weights. Highly ranked sentences are extracted from the input document so it extracts important sentences which directs to a high-quality summary of the input document. According to this research, there are basically two methods of text summarization. First one is Extractive summarization and the second one is Abstractive summarization. Extractive summarization is basically creating a summary based on strictly what you get in the text. It can be compared to copying down the main points of a text without any modification to those points and rearranging the order of that points grammar to make more sense out of Abstractive summarization techniques tend to mimic the process of 'paraphrasing' from a text than just simply summarizing it. Texts summarized using this technique looks more human-like and produce more condensed summaries.

In the proposed approach of this paper, researchers use extractive method to get summary of a given output. Extractive text summarization is divided into two phases as Pre-processing and Processing. They have divided this pre-processing and processing stages into a few steps. Firstly, the file which is given as input is tokenized in order to get tokens of the terms. Then the stop words are removed from the text after tokenization. The words which remained are considered as key words. Then key words are taken as an input and a part of tag is attached to each key word. After completing this pre-processing stage, frequency of each keyword is calculated. As an example, they calculate how frequently a particular key word has occurred and from this, maximum frequency of the keyword is taken. Then, the weighted frequency if the word is calculated by dividing frequency of the key words by maximum frequency of the key words. In this step, the sum of weighted frequencies is calculated. Finally, summarizer extracts the high weighted frequency sentences and the extracted sentences are converted into audio form. The research concludes that the proposed model improves the accuracy when compared to traditional approaches.

1.2. Research Gap

According to the available research papers and resources, there are several systems developed related to smart learning. So far, there is no simple, cohesive concept of smart learning. The definition of smart learning is constantly discussed by multidisciplinary scholars and educational professionals. In general, the atmosphere for intelligent learning is productive, efficient and engaging. In this section, the research gap between previous researches/products and the Note Taker component of the EduEasy Smart Learning Assistant System is discussed.

The Note Taker is one of the four main components of the EduEasy Smart Learning Assistant System. This component is designed to do two particular tasks. As the first task, the program converts the lecturers voice in to text and create a voice transcript. Then the second task is creating a summarized note. The program takes the previously created voice transcript and creates a summarized note.

As stated in the previous section, there are many researches done on voice transcriptions [6]-[7] and text summarization [8]-[10]. In those speech and lecture transcription researches, they have used different and novel methods to transcribe audio into text. Also, in those text summarization researches, they have used different type of methods and algorithms to do summarizations but a comparison of effectiveness between those methods is a significant gap. Some researchers have done some calculations to determines the accuracy of the effectiveness of their methods but despite the numbers, a summary should be effective to a person reading that. So the opinion of readers matters the most but researching that area has not been done before.

There are online word processing tools (i.e. LinguaKit) to summarize texts and do speech-to-text transcriptions. But there is no research or product which is a combination of voice transcription and text summarization. So in this Note Taker component of the EduEasy Smart Learning Assistant System, those two features are combined and a full functioning speech-to-text converter and a text summarizer is implemented.

2. RESEARCH PROBLEM

Students who study smart instead of study hard, can have good success in their education. But studying smart is a challenge without a correct support. University students get used to study their lessons in various ways. EduEasy is a smart learning assistant system for students to effectively learn and revise lectures done at the university. The Note Taker is one of the main four components of the EduEasy Smart Learning Assistant System which is designed to serve a specific purpose.

Most of the time university students do not have a proper lecture note for various reasons. Some students do not attend lectures regularly or even if they do attend lectures, they may not write down important points etc. Students who do not have proper notes may feel uncomfortable during exam periods because without a good note, it is very difficult to understand theories and other content. Then they have to waste their time finding notes instead of studying. The Note Taker component is designed to provide a solution for this problem. This application generates a summarized note using a transcription of the lecturer's voice. This speech-to-text converting facility is also provided by the application. After that, students are able to revise the notes anytime they want.

3. RESEARCH OBJECTIVES

3.1. Main Objective

• Develop an E-leaning application for students to effectively learn and revise lectures done at the university.

3.2. Specific Objectives

- Capture the lecturer's voice using microphone and voice recorder.
- Develop a speech-to-text converter to transcribe lectures into text.
- Develop a text summarizer to summarize the transcription.
- Conduct a research/survey to find out the most effective text summarization algorithm among popular algorithms.
- Develop an attractive and easy to use user interface for the speech-to-text converter and the text summarizer.
- Deploy the web application.

4. METHODOLOGY

4.1. Methodology

Students who study smart instead of study hard, can have good success in their education. But studying smartly is a challenge without the correct support. EduEasy is an E-learning application for students to effectively learn and revise lectures done at university. EduEasy application has mainly four components. They are note taker, reference finder, question presenter and slide matcher. Each of those components will help to overcome problems which university students has to face on a daily basis.

University students get used to study their lessons in various ways. Students who do not go lectures regularly may not have good lecture notes. Students who don't have proper notes, may feel uncomfortable during exam periods because without having a good note it is very difficult to understand the theories and learn them. Then they waste their time to find out good notes. EduEasy application is designed to automatically generate a summarized short note by first, recording the lecturer's voice & converting it into a text document and then, using a text summarization algorithm, removing unnecessary information & converting the text document into a summarized note.

The Note Taker component of the EduEasy application is divided into two basic parts. First one is Speech-to-Text converter and the second one is the Text Summarizer. Speech-to-Text converter is designed to convert lecturers voice to text. This process can be explained using a few steps.

First, the lecturer's voice should input through the microphone. This can be done by either real time or recording the voice to another device and play it to the microphone later. Reason for using this method instead of converting an entire lecture recording audio file at once is that, converting large audio file takes a long time than converting real time. So to increase the efficiency of the application, this method was used.

Then the audio is converted into text using JavaScript Web Speech API. The Web Speech API incorporates voice data into web applications. The Web Speech API has two main parts, SpeechRecognition (Speech-to-Text) and SpeechSynthesis (Text-to-Speech). To develop the Speech-to-Text converter function in the Note Taker

component, speech recognition is used. Speech recognition is accessed via the SpeechRecognition interface, which provides the ability to recognize voice context from an audio input (normally via the device's default speech recognition service) and respond appropriately. Generally, the interface's constructor is used to create a new SpeechRecognition object, which has a number of event handlers available for detecting when speech is input through the device's microphone. Speech synthesis is accessed via the SpeechSynthesis interface, a text-to-speech component that allows programs to read out their text context (normally via the device's default speech synthesizer).

Second part of the Note Taker component is the Text Summarizer. This is a tool to analyze text documents and simplify them using keywords. It captures most significant concepts, providing a concise overview which can be useful in studying and memorizing for students. The text summarization process is developed based on the Term Frequency - Inverse Document Frequency (TF-IDF / TF*IDF) algorithm. TF-IDF is an information retrieval technique that weights a term frequency and its inverse document frequency. Each word or term has its respective TF and IDF score. The product of the TF and IDF scores of a term is called TF-IDF weight of that term. A high weight in TF-IDF is reached by a high term frequency in the given document and a low document frequency of the term in the whole collection of documents.

TF-IDF algorithm has many uses, most importantly in automated text analysis, and it is very useful for scoring words in machine learning algorithms for Natural Language Processing (NLP). Once the words are transformed into numbers, in a way that's machine learning algorithms can understand, the TF-IDF score can be fed to algorithms such as Naive Bayes and Support Vector Machines, greatly improving the results of more basic methods like word counts. TF-IDF algorithm works by increasing proportionally to the number of times a word appears in a document, but is offset by the number of documents that contain the word. So, words that are common in every document, such as this, what, and if, rank low even though they may appear many times, since they don't mean much to that document in particular.

TF-IDF algorithm is made of two algorithms multiplied together.

a) Term Frequency (TF)

Term frequency is how often a word appears in a document, divided by how many words there are.

$$TF(t) = \frac{Number of times (t) appear in a document}{Total number of terms in the document}$$

b) Inverse Document Frequency (IDF)

Term frequency is how common a word is, but inverse document frequency is finding out how unique or rare a word is.

$$TF(t) = \frac{Number of times (t)appear in a document}{Total number of terms in the document}$$

To choose this particular method for development, a survey has been conducted. Researchers have done calculation to find out the accuracy of these algorithms. But according to the EduEasy research standpoint, opinions of the readers matter the most because a summary is generated for human understanding and a reader should understand and appreciate a summary. So, two summaries of a same text are generated using the TF-IDF algorithm & another popular text summarization method, Sentence Scoring algorithm and the results were published online for the effectiveness review survey. Majority of the participants chose the summary generated using the TF-IDF algorithm as the most effective option. Results of this survey can be reviewed in the "Results" section.

Following diagram (Figure 1) is a simplified illustration of all the NLP processes in the Note Taker component.

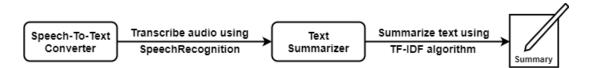


Figure 1: NLP Processes of the Note Taker component

4.1.1. System Architecture

The Figure 2 depicts how the Note Taker component works and the process begins with the note taker component by recording the lecturer's voice to the computer. Then using a transcription algorithm (SpeechRecognition interface of Web Speech API), audio is converted into text format. This ends the audio transcription part of the component. Then the summarizing part begins. The transcribed text is converted into a summarized note by the text summarizer developed using Python 3.6 & TF-IDF text summarization algorithm. To choose that particular algorithm, an online survey had been conducted and the majority of people preferred the TF-IDF algorithm's summary over Sentence Scoring algorithm's summary of the same text. Then the summarized note opens and the lecturer can submit the note. After that, students can refer to it from their computers.

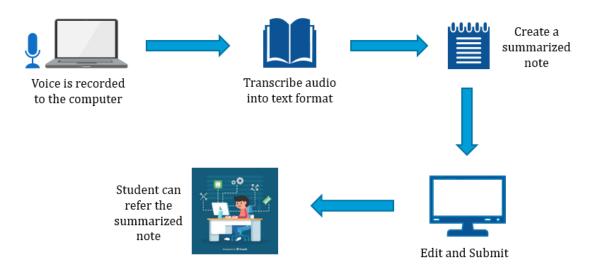


Figure 2: System Diagram of the Note Taker

4.1.2. User Interface Implementation

Following Figure 3, Figure 4 & Figure 5 depicts the three main user interfaces of the Note Taker component, Speech-to-Text converter, text summarizer and summarized note interface.

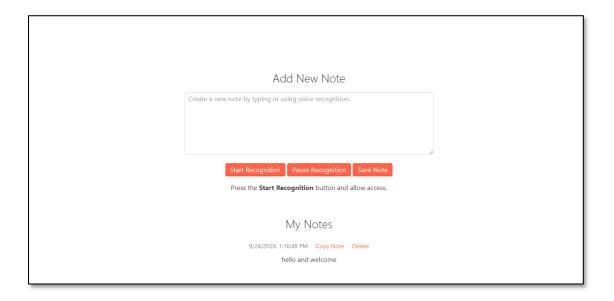


Figure 3: Speech-to-Text Converter UI

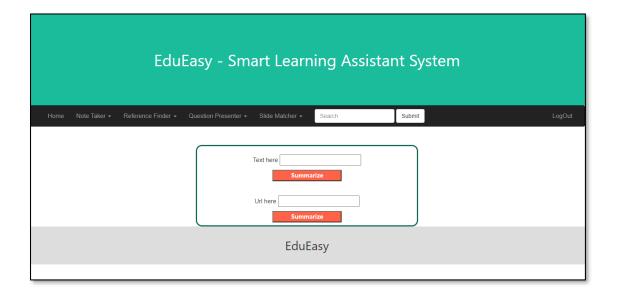


Figure 4: Text Summarizer UI

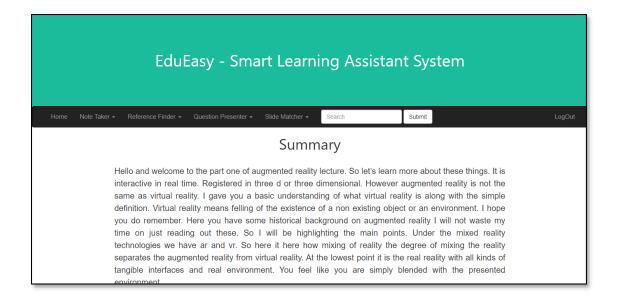


Figure 5: Summary Display UI

4.1.3. Project Management

Project management is a fundamental concept that can be put into effect with every project undertaken. This concept is more important to broader tasks carried out by teams and can contribute to project performance by adding project management guidelines to a project. When designing the web application "EduEasy – Smart Learning Assistant System," some of the concepts of project management were implemented in this project in order to achieve a consistent direction that could be traced from the beginning of the project to the end so that the created application fulfills the predefined goals and objectives. Software Development Life Cycle (SDLC) is a structured software construction mechanism that guarantees the consistency and correctness of the developed apps. The goal of the SDLC process is to deliver high quality software that meets the customers' expectations. In the present time frame and cost the system implementation should be complete. SDLC consists of a comprehensive strategy that describes how to develop, create, and manage complex applications. This phase of the SDLC's life cycle has its own mechanism and deliverables that feed into the next phase. SDLC stands for software development.



Figure 6: Project Management

Identification and Initiation - The complexity of the project was closely analyzed during the project implementation process in order to identify whether the project could be reasonably completed within the given time period. In dealing with the exact project scale, time constraints became a big concern. Depending on the utility to the customers and depending on the time limit, the features of the programs had to be chosen. In

addition, before beginning to prepare the scheme, the feasibility aspect was also considered at this initial stage. In order to determine the best practices available from the current implementations, numerous research have been performed.

Definition and Planning - All the required tasks had to be arranged according to a clear timeline during the project planning period to meet the requirements of the project within the specified timeframe. Even the necessary tools for the project were analyzed during the planning period. Since the project was designed to be carried out as a web application, before beginning the production process, the necessary technology and languages had to be researched and taught. The most suitable tools and software were chosen after review of different tools, considering the appropriateness of such tools and software to the project.

After reviewing many projects and strategies that were pursued within them, it was found that most of the projects that were performed in compliance with a 10 predefined schedule and a time frame were able to effectively complete the project without any difficulties Getting a clear plan often decreased project costs when evaluating commercial projects that are performed by multiple teams. The functional requirements of the project were to be based mainly on the specified time period for implementation. All the roles that had to be created were split down into subtasks with set deadlines for this reason.

Development and Launch - During this deployment process, the planned functions and the pre-considered technology were combined to arrive at a final deliverable. The "EduEasy - Smart Learning Assistant Software" was difficult to create with time limits, and due to the implementation of new technologies not seen in previous programs. For the implementation scenario an approach such as waterfall or iterative waterfall was not feasible. Therefore, agile approach considering multiple variables was adopted. Agile method has become more practical to use than waterfall or iterative method, because checking can be done by using agile method after each subtask of the software has been developed. With the production process, training can be carried out that helps the team to be versatile during training in modifying multiple parts, making it easier for the team. Using agile approach also helps the team to look ahead of the other activities

to be developed and planned for the implementation process while planning is being carried out and software is being carried out for the project.

In developing the "EduEasy - Smart Learning Assistant System", test driven development was the best approach to be used due the flexibility it provided while developing to meet the exact requirement within the given time constraint. As one subtask could be fully designed, evaluated, iterated until the testing passed and then transferred to the next subtask, the implementation process was more feasible.

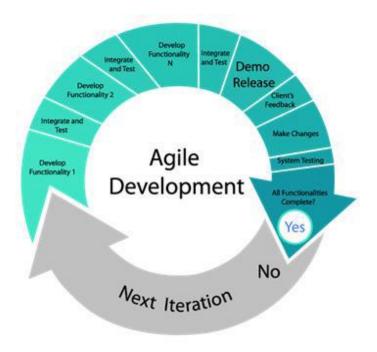


Figure 7: Agile Development

Closeout - This process is largely about launching the final deliverable. The performance of the method may be defined in this stage according to the input provided by the application's users. In order to assess the lessons learnt from the production of the application and the possible improvements that are to be implemented in order to increase the product quality, this stage is so important.

4.1.4. Development

As the backend development programming language of the application, Python has been used as it is more productive language and an interpretive language which is accompanied by elegant syntaxes and libraries. Python is an excellent language choice for scripting and rapid application development in many areas. It is a dynamically typed programming language. For this specific development process, Python version 3.5 was used.



Figure 8: Python Programming Language

Choosing the necessary tools for system development plays a significant role for the success of a program or an application. As this application was designed as a web application and Python 3.5 was used as the programming language, it was decided to use Anaconda Spyder IDE as the development IDE. Spyder IDE is a powerful scientific environment written in Python, for Python. If features unique combination of the advanced editing, analysis, debugging and profiling functionality of a comprehensive development too with beautiful visualization capabilities.



Figure 9: Anaconda Spyder IDE

To develop the front-end of the component and connect it to the back-end, Flask framework was used. Flask is a micro web framework written in Python. It is classified as a micro framework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions.



Figure 10: Flask Framework

4.1.5. Software System Attributes

Usability - This would be a strong sign of incorrect goals. The first goal of usability is performance and efficacy, while esthetic meaning comes after the commodity has been proved to be functional. Usability is human-centered, capturing and knowing the entire range of consumer desires is easier than only a collection of diverse data. End-users are the greatest source of details, but the more nuanced products and websites have many user classes, so it is much more important to get all the details, around the board, not just desires but wants and expectations, the complete range of details allows to maintain continuity and authenticity before finalizing the course of usability. Usability innovation can be used to define the needs of consumers for customer retention and performance. In order to meet the needs of the targeted clients according to their criteria, measures must be taken in order to address the needs of the clients and to define their key objectives.

Reliability - In order for an application to be accurate, it should always run without any faults or errors in the stated environment for a given time span. Tech Stability implies operational efficiency. It is defined as the ability of a device or part to perform its necessary functions under static conditions for a specified period of time. Software Reliability is a core component of software quality, consisting of features, accessibility, consistency, serviceability, ability to update, maintainability, and documentation.

Scalability - Software scalability is the feature of a method or device to expand its capability and flexibility depending on the demand of its users. Scalable applications will stay robust when adapting to updates, improvements, modifications, and resource reduction. In a fast-moving market, you need to develop your products with development in mind. Software scalability is important for the survival of your enterprise. Read on to see if you can make the apps scalable. This is where the scalability of software comes in. Your commodity should be able to fulfill its purpose from now on and can also be modified on the basis of consumer demand. In short, with development in mind, you need to develop your product.

Maintainability - Technology will still require additional functionality or bug fixes. Maintainable software is easy to expand and repair, which enables the uptake and use of software. We will guide you on the design and production of sustainable applications that can help both you and your users. It would be better practice to provide specific names and descriptions in the forms, which would make it simpler later if any improvements were required. As a method for the success of this application, agile principles have been followed. When beginning a project with a simple architecture, it will be easier to extend later without making it more difficult to lead developers in trouble by adding additional features.

There are a variety of explanations for retaining the program after you have delivered it to the customer:

- Bug fixing It requires looking for and fixing errors to enable the program to run smoothly
- Capability enhancement Enhancement of software to include the latest features needed by customers

- Replacement -replace unnecessary functionality to increase addictiveness and efficiency
- Security Issues Fix security vulnerabilities found in your proprietary code or third party code, particularly open source components.

4.1.6. Operations

"EduEasy Smart Leaning Assistant System" web application and the Note Taker component of the system allows users to do following operations.

- User can register to the application.
- User can log in to the application using credentials.
- Users can click "Stat Recording" button in the Speech-to-Text converter and record audio.
- Users can save automatically converted text notes.
- Users can listen to saved notes.
- Users can generate a summarized version of the saved notes.
- Users can generate summaries using any text or by providing a link of a note anywhere from the internet.

4.1.7. User Characteristics

"EduEasy - Smart Learning Assistant System" web application mainly targets for university students as well as university lecturers. They are the most suitable group of people to use and experience the benefits of the application.

4.2. Commercialization Aspects of the Product

EduEasy E-learning web application is mainly designed for university environment. Students who wants to study smart, EduEasy system will help them to make their works easy. Students are struggling near exam to find a proper lecture notes, and references, sample questions and the course materials such as lecture PPTs. EduEasy application is a combination of those all requirements.

EduEasy application is introduced with mainly three packages: Platinum, Gold and Silver. All the functionalities are included to the platinum package. Gold package includes all functions except reference finding function. Silver package has only the lecture summarization function. Furthermore, a one-month free trial will be given to all new users of EduEasy system.

Features provided by the EduEasy system

- A summarized note of the lecture conducted by the lecturer for each day.
- Generate automatically the questions and answers for the summarized note.
- Generate automatically the references for additional reading.
- Navigate to the exact lecture slide from summarized note.

Benefits of the EduEasy system

- Easy to find the summarized note.
- Easy to find similar questions and answers.
- Can find relevant references easily.
- Can navigate to the relevant lecture slide automatically.

5. TESTING & IMPLEMENTATION RESULTS & DISCUSSION

5.1. Testing

Testing is the method of evaluating a system or component(s) in order to decide whether or not it meets the specifications defined. Simply put, testing is conducting a method to find any holes, defects or incomplete specifications as opposed to the actual specifications.

And like this, testers too will begin testing. The cost and time to rework and create error-free software that is delivered to the customer is minimized by an early start to testing. However, testing will start from the Requirements Gathering stage in the Software Development Life Cycle (SDLC) and continue until the software is deployed.

When to stop testing is difficult to decide, since testing is a never-ending process and no one can say that a program is 100% tested. To stop the research process, the following factors must be taken into process.

- Testing Deadlines
- Completion of test case execution
- Completion of functional and code coverage to a certain point
- Bug rate falls below a certain level and no high-priority bugs are identified
- Management decision

During the process of testing, there are various stages. A short overview of these stages is given here. Test thresholds provide various methodologies which can be used during software development. The key software testing levels are

- Functional Testing
- Non-functional Testing

Functional Testing - This is a method of black-box testing that is focused on the program requirements that are to be evaluated. The program is checked by supplying input, and the results that need to adhere to the features for which it was intended are then analyzed. Functional program testing is carried out on a complete, integrated system to assess conformity of the system with the defined specifications. Functional

testing is a much-needed testing technique for any application as it ensures that the functionality in the application have been developed in 21 accordance with the functions that have been specified. In the "EduEasy-Smart Learning Assistant System" web application, acceptance testing was carried out in order to test for functionality.

Non-functional Testing - Checking an application from its non-functional attributes is based on this section. Non-functional testing includes evaluating applications based on requirements that are non-functional but essential in nature, such as performance, security, user interface, etc.

To test the functionalities of the text summarizer part of the Note Taker component, following test cases was used.

Test Case 01

"Those Who Are Resilient Stay In The Game Longer

On the mountains of truth, you can never climb in vain: either you will reach a point higher up today, or you will be training your powers so that you will be able to climb higher tomorrow. Friedrich Nietzsche.

Challenges and setbacks are not meant to defeat you, but promote you. However, I realise after many years of defeats, it can crush your spirit and it is easier to give up than risk further setbacks and disappointments. Have you experienced this before? To be honest, I don't have the answers. I can't tell you what the right course of action is; only you will know. However, it's important not to be discouraged by failure when pursuing a goal or a dream, since failure itself means different things to different people. To a person with a Fixed Mindset failure is a blow to their self-esteem, yet to a person with a Growth Mindset, it's an opportunity to improve and find new ways to overcome their obstacles. Same failure, yet different responses. Who is right and who is wrong? Neither. Each person has a different mindset that decides their outcome. Those who are resilient stay in the game longer and draw on their inner means to succeed. I've coached many clients who gave up after many years toiling away at their respective goal or dream. It was at that point their biggest breakthrough came. Perhaps all those years of perseverance finally paid off. It was the 19th Century's minister Henry Ward Beecher who once said: "One's best success comes after their greatest disappointments." No one knows what the future holds, so your only guide is whether you can endure repeated defeats and disappointments and still pursue your dream. Consider the advice from the American academic and psychologist Angela Duckworth who writes in Grit: The Power of Passion and Perseverance: "Many of us, it seems, quit what we start far too early and far too often. Even more than the effort a gritty person puts in on a single day, what matters is that they wake up the next day, and the next, ready to get on that treadmill and keep going. I know one thing for certain: don't settle for less than what you're capable of, but strive for something bigger. Some of you reading this might identify with this message because it resonates with you on a deeper level. For others, at the end of their tether the message might be nothing more than a trivial pep talk. What I wish to convey irrespective of where you are in your journey is: NEVER settle for less. If you settle for less, you will receive less than you deserve and convince yourself you are justified to receive it.

I recall a passage my father often used growing up in 1990s: "Don't tell me your problems unless you've spent weeks trying to solve them yourself. That advice has echoed in my mind for decades and became my motivator. Don't leave it to other people or outside circumstances to motivate you because you will be let down every time. It must come from within you. Gnaw away at your problems until you solve

them or find a solution. Problems are not stop signs, they are advising you that more work is required to overcome them. Most times, problems help you gain a skill or develop the resources to succeed later. So embrace your challenges and develop the grit to push past them instead of retreat in resignation. Where are you settling in your life right now? Could you be you playing for bigger stakes than you are? Are you willing to play bigger even if it means repeated failures and setbacks? You should ask yourself these questions to decide whether you're willing to put yourself on the line or settle for less. And that's fine if you're content to receive less, as long as you're not regretful later. If you have not achieved the success you deserve and are considering giving up, will you regret it in a few years or decades from now? Only you can answer that, but you should carve out time to discover your motivation for pursuing your goals. It's a fact, if you don't know what you want you'll get what life hands you and it may not be in your best interest, affirms author Larry Weidel: "Winners know that if you don't figure out what you want, you'll get whatever life hands you." The key is to develop a powerful vision of what you want and hold that image in your mind. Nurture it daily and give it life by taking purposeful action towards it. Vision + desire + dedication + patience + daily action leads to astonishing success. Are you willing to commit to this way of life or jump ship at the first sign of failure? I'm amused when I read questions written by millennials on Quora who ask how they can become rich and famous or the next Elon Musk. Success is a fickle and long game with highs and lows. Similarly, there are no assurances even if you're an overnight sensation, to sustain it for long, particularly if you don't have the mental and emotional means to endure it. This means you must rely on the one true constant in your favour: your personal development. The more you grow, the more you gain in terms of financial resources, status, success. If you leave it to outside conditions to dictate your circumstances, you are rolling the dice on your future. So become intentional on what you want out of life. Commit to it. Nurture your dreams. Focus on your development and if you want to give up, know what's involved before you take the plunge. Because I assure you, someone out there right now is working harder than you, reading more books, sleeping less and sacrificing all they have to realise their dreams and it may contest with yours. Don't leave your dreams to chance."

Test Case 02

"Hello and welcome to the part one of augmented reality lecture. So today here we are going to talk about augmented reality and I do believe right now you are not you all are not strangers to this word augmented reality or virtual reality. So you can see by the given images you can see here blended all these three images they show some cg or computer graphics rendered with the real images or real photographs that is coming from a camera. So let's learn more about these things. So basically augmented reality is a field of computer research which deals with the combination of real worlds and computer generated data or in this case computer generated imagery. Ronald a famous scientist defines an augmented reality system as one that overlay of computer presented information on top of the real world. If you see the previous image these are the overlaid computer generated images or text on the real word. It combines real and virtual realities. It is interactive in real time. Registered in three d or three dimensional. However augmented reality is not the same as virtual reality. I gave you a basic understanding of what virtual reality is along with the simple definition. Virtual reality means felling of the existence of a non existing object or an environment. I hope you do remember. Here you have some historical background on augmented reality I will not waste my time on just reading out these. So I will be highlighting the main points. Mixing of realities. Even though some time back we have two different word like virtual reality and augmented reality right now the whole thing became into one particular domain that is called mixed reality. Under the mixed reality technologies we have ar and vr. So here it here how mixing of reality the degree of mixing the reality separates the augmented reality from virtual reality. Or how mixing of reality separates the reality from computer generated reality. So when the computer generated basically is low it is simply the real reality. At the lowest point it is the real reality with all kinds of tangible interfaces and real environment. Little by little when you introduce the computer generated imagery like optical see through or some overlaying computer generated content it becomes augmented reality an it becomes video reality and finally when you have the whole environment developed using the computer generated components we brings out a virtual environment and with that when you have the immerse experience regardless of its real or unreal environment which is generated using computer generated imagery then that is called the virtual reality. Because you simply fell that you are in a real environment regardless of that computer generated components seem real or not. You feel like you are simply blended with the presented environment."

Test Case 03

"A thesis or dissertation[1] is a document submitted in support of candidature for an academic degree or professional qualification presenting the author's research and findings.[2] In some contexts, the word "thesis" or a cognate is used for part of a bachelor's or master's course, while "dissertation" is normally applied to a doctorate, while in other contexts, the reverse is true.[3] The term graduate thesis is sometimes used to refer to both master's theses and doctoral dissertations.[4] The required complexity or quality of research of a thesis or dissertation can vary by country, university, or program, and the required minimum study period may thus vary significantly in duration. The word "dissertation" can at times be used to describe a treatise without relation to obtaining an academic degree. The term "thesis" is also used to refer to the general claim of an essay or similar work. The term "thesis" comes from the Greek θέσις, meaning "something put forth", and refers to an intellectual proposition. "Dissertation" comes from the Latin dissertātiō, meaning "discussion". Aristotle was the first philosopher to define the term thesis. "A 'thesis' is a supposition of some eminent philosopher that conflicts with the general opinion...for to take notice when any ordinary person expresses views contrary to men's usual opinions would be silly".[5] For Aristotle, a thesis would therefore be a supposition that is stated in contradiction with general opinion or express disagreement with other philosophers (104b33-35). A supposition is a statement or opinion that may or may not be true depending on the evidence and/or proof that is offered (152b32). The purpose of the dissertation is thus to outline the proofs of why the author disagrees with other philosophers or the general opinion. A thesis (or dissertation) may be arranged as a thesis by publication or a monograph, with or without appended papers, respectively, though many graduate programs allow candidates to submit a curated collection of published papers. An ordinary monograph has a title page, an abstract, a table of contents, comprising the various chapters like introduction, literature review, methodology, results, discussion, and bibliography or more usually a references section. They differ in their structure in accordance with the many different areas of study (arts, humanities, social sciences, technology, sciences, etc.) and the differences between them. In a thesis by publication, the chapters constitute an introductory and comprehensive review of the appended published and unpublished article documents. Dissertations normally report on a research project or study, or an extended analysis of a topic. The structure of a thesis or dissertation explains the purpose, the previous research literature impinging on the topic of the study, the methods used, and the findings of the project. Most world universities use a multiple chapter format: a) an introduction: which introduces the research topic, the methodology, as well as its scope and significance b) a literature review: reviewing relevant literature and showing how this has informed the research issue c) a methodology chapter, explaining how the research has been designed and why the research methods/population/data collection and analysis being used have been chosen d) a findings chapter: outlining the findings of the research itself e) an analysis and discussion chapter: analysing the findings and discussing them in the context of the literature review (this chapter is often divided into two—analysis and discussion) f) a conclusion.:[6][7] which shows judgement or decision reached by thesis"

5.2. Results

The Speech-to-Text converter was developed in the first half of the note taker process. Both the options of converting a recorded voice clip to text and converting voice to text real time was considered. During the testing, an observation was, converting a long voice record of a lecture to text takes a lot of time to process using the SpeechRecognition API. But when converting speech-to-text real time, processing time is significantly low because it takes a single word or a phrase at a time instead of trying to covert a whole audio file at once. So it was determined that the real time converting method is more suitable because it is time efficient. In the developed speech-to-text converter, Word Error Rate (WER) of the transcriptions are around 30-35% which is better than the normal rate 40-45% [7]. Then the Text Summarizer was developed and a research survey has been conducted to determine the best abstractive text summarizing algorithm for the text summarization process. As the result of that survey (Figure 11), the TF-IDF algorithm was selected as the most effective algorithm with the choice of 80% respondents. Connected the above two sub-components & tested the note taker using a lecture recording as input. The output summarized note is nearly 1/4 long as the full voice transcription.

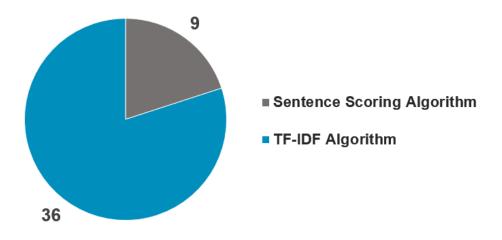


Figure 11: Results of the survey

The text summarizer was developed using the TF-IDF algorithms. Summarized results of the previously stated test cases are shown below.

Test Result 01

"Have you experienced this before? To be honest, I don't have the answers. Same failure, yet different responses. Who is right and who is wrong? Neither. Each person has a different mindset that decides their outcome. It was at that point their biggest breakthrough came. Perhaps all those years of perseverance finally paid off. It must come from within you. Gnaw away at your problems until you solve them or find a solution. Where are you settling in your life right now? Could you be you playing for bigger stakes than you are? Success is a fickle and long game with highs and lows. Commit to it. Nurture your dreams. Don't leave your dreams to chance.

Have you experienced this before? To be honest, I don't have the answers. I can't tell you what the right course of action is; only you will know. Same failure, yet different responses. Who is right and who is wrong? Neither. Each person has a different mindset that decides their outcome. It was at that point their biggest breakthrough came. Perhaps all those years of perseverance finally paid off. Some of you reading this might identify with this message because it resonates with you on a deeper level. For others, at the end of their tether the message might be nothing more than a trivial pep talk. What I wish to convey irrespective of where you are in your journey is: NEVER settle for less. It must come from within you. Gnaw away at your problems until you solve them or find a solution. Problems are not stop signs, they are advising you that more work is required to overcome them. So embrace your challenges and develop the grit to push past them instead of retreat in resignation. Where are you settling in your life right now? Could you be you playing for bigger stakes than you are? Are you willing to play bigger even if it means repeated failures and setbacks? Are you willing to commit to this way of life or jump ship at the first sign of failure? Success is a fickle and long game with highs and lows. The more you grow, the more you gain in terms of financial resources, status, success. Commit to it. Nurture your dreams. Don't leave your dreams to chance"

Test Result 02

"Hello and welcome to the part one of augmented reality lecture. So let's learn more about these things. It is interactive in real time. Registered in three d or three dimensional. However augmented reality is not the same as virtual reality. I gave you a basic understanding of what virtual reality is along with the simple definition. Virtual reality means felling of the existence of a non existing object or an environment. I hope you do remember. Here you have some historical background on augmented reality I will not waste my time on just reading out these. So I will be highlighting the main points. Under the mixed reality technologies we have ar and vr. So here it here how mixing of reality the degree of mixing the reality separates the augmented reality from virtual reality. At the lowest point it is the real reality with all kinds of tangible interfaces and real environment. You feel like you are simply blended with the presented environment."

Test Result 03

"The word "dissertation" can at times be used to describe a treatise without relation to obtaining an academic degree. The term "thesis" is also used to refer to the general claim of an essay or similar work. The term "thesis" comes from the Greek θέσις, meaning "something put forth", and refers to an intellectual proposition. "Dissertation" comes from the Latin dissertātiō, meaning "discussion". Aristotle was the first philosopher to define the term thesis. A supposition is a statement or opinion that may or

may not be true depending on the evidence and/or proof that is offered (152b32). The purpose of the dissertation is thus to outline the proofs of why the author disagrees with other philosophers or the general opinion. They differ in their structure in accordance with the many different areas of study (arts, humanities, social sciences, technology, sciences, etc.) and the differences between them. In a thesis by publication, the chapters constitute an introductory and comprehensive review of the appended published and unpublished article documents. Dissertations normally report on a research project or study, or an extended analysis of a topic. :[6][7] which shows judgement or decision reached by thesis⁹⁹

5.3. Research Findings

The main outcome of this application is to develop an e-learning application for students to effectively learn and revise lectures done at the university. The Note Taker component is developed as the speech-to-text converter & the text summarizer tool of the application. Following are research findings related to the Note Taker component.

- Converting a long voice record into text takes a lot of time to process than converting speech-to-text real time using the Web Speech API.
- Word Error Rate (WER) is higher when converting speech-to-text real than converting the whole voice record at once.
- 80% of the respondents of the survey selected the summary of TF-IDF algorithm rather than the summary of sentence scoring algorithm, so as conclusion, TF-IDF algorithm is a more effective text summarizing algorithm.

5.4. Discussion

- When the Web Speech API tries to convert a long voice recording into text, it takes a significantly long time to process because it tries to process the whole audio file at once. But when converting speech-to-text real time, API converts words or phrases at a time. So it takes a lower amount of time to process.
- Accuracy of real time speech-to-text conversion is significantly low because the program tries to identify audio wave through a microphone and transcribe at the same time. Even the background noises in the environment can affect the outcome. But when converting an audio file directly, the program does not need to listen through microphones. So from the accuracy standpoint, converting an audio file is better than converting speech-to-text real time.

6. CONCLUSION

The project "EduEasy - Smart Learning Assistant System" is a E-learning web application developed using a combination of several latest technologies which helps university students to overcome some of their daily problems and effectively learn and revise lectures done at the university. The main focus of the application is to help students to make summarize notes, find online references, generate lecture related questions and matching the notes and lecture slides for ease of use. To achieve that goal, applications has four different components which are The Note Taker, The Reference Finder, The Question Presenter and The Slide Matcher.

The Note Taker is one of the main four components of the EduEasy Smart Learning Assistant System which is designed to serve a specific purpose. Most of the time university students do not have a proper lecture note for various reasons. Students who do not have proper notes may feel uncomfortable during exam periods because without a good note, it is very difficult to understand theories and other content. The Note Taker component is designed to provide a solution for this problem. This application generates a summarized note using a transcription of the lecturer's voice. This speech-to-text converting facility is also provided by the application. After that, students are able to revise the notes anytime they want.

The application is still in the developing and improving stage. Future goal is to deploy the application for active usage and beta testing with government and private universities with the help of Ministry of Education. Then the application will be released for commercial use with different levels of functionalities.

7. REFERENCES

- [1] J. Meer, "Students' note-taking challenges in the twenty-first century: Considerations for teachers and academic staff developers," in Teaching *in Higher Education*, 17, February, 2012, pp. 13-23.
- [2] M. A. Awar, "Pioneering smar learning". [Online]. Available: https://www.ellucian.com/emea-ap/insights/pioneering-smart-learning. [Accessed May 4, 2020].
- [3] A. Jokiaho, B. May, M. Specht and S. Stoyanov, "Barriers to suing E-leaning in an Advanced Way," in *International Journal of Advanced Corporate Learning*, 2018, vol. 11, no. 1, pp. 17-22.
- [4] D. Kaur, "How Smart Class Techology is Benefiting Education Sector". Available: https://www.entrepreneur.com/article/322587. [Accessed May 4, 2020].
- [5] CAE Team, "What is Smart Learning and why does it interest educational centers?" cae.net, 2020. [Online]. Available: https://www.cae.net/what-is-smart-learning-and-why-does-it-interest-educational-centers/. [Accessed: Feb. 21, 2020].
- [6] T. Kawahara, H.Nanjo and S. Furui, "Automatic transcription of spontaneous lecture speech," in *IEEE Workshop on Automatic Speech Recognition and Understanding*, 2001, ASRU '01., Madonna di Campiglio, Italy, 2001, pp. 186-189.
- [7] C. Munteanu, G. Penn, and R. Baecker, "Web-Based Language Modeling for Automatic Lecture Transcription," in 8th Annual Conference of the International Speech Communication Association, Antwerp, Belgium, August, 2007, pp. 2353-2356.
- [8] K. S. Thakkar, R. V. Dharaskar, and M. B. Chandak, "Graph-Based Algorithms for Text Summarization," in *3rd International Conference on Emerging Trends in Engineering and Technology, Goa, India*, 2010, pp. 516-519.
- [9] N. K. Nagwani, S. Verma, "A Frequent Term and Semantic Similarity based Single Document Text Summarization Algorithm," in *International Journal of Computer Applications*, March, 2011, vol. 17, no. 2.
- [10] J. N. Madhuri, R. Ganesh Kumar, "Extractive Text Summarization Using Sentence Ranking," in 2019 International Conference on Data Science and Communication (IconDSC), Bangalore, India, 2019, pp. 1-3.

8. APPENDICES

Appendix A

Following Figure 12 is the Gantt Chart of the application development timeline.

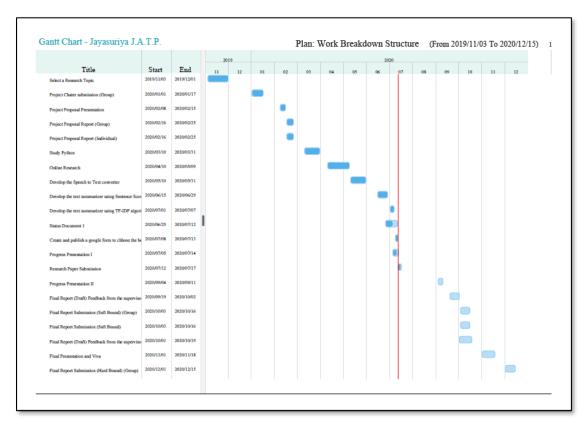


Figure 12: Gantt Chart