International Journal of Research in Science and Technology
Volume 12, Issue 01: January - March 2025
ISSN 2394 - 9554

Transcription and Text Summarization using Deep Learning Techniques

R Manimegalai *, DeepakRaj R K Inderjith K Rajesh V Susmeta A

Department of Computer Science and Engineering, PSG Institute of Technology and Applied Research Coimbatore, Tamil Nadu, India

* drrm@psgitech.ac.in

Abstract
Large amount of data are generated every day in the Internet. There is a need for more efficient techniques for effectively obtaining crucial information from the large data generated. One method for locating the most important and useful information in a document is text summarization. It condenses it into a shorter form while preserving the primary idea. Voice-powered transcription reduces the time by taking care of the preliminary transcribing in real-time. The efficiency is improved as the generated summary has proper context and the most important information. Voice recognition techniques help in generating the document quickly. The system is designed to use audio inputs from the user to convert it into textual format. This text along with a compression ratio is provided as an input to a model that summarizes the text using a combination of the algorithms. Sentences are graded based on the weights. The most important lines that maintain the critical information are generated as the text summary. The model used for summarization is made up of multiple sub-modules that assign the scores for all the sentences in an exclusive manner. These scores are later aggregated to find the most important lines which are then displayed to the user.

Keywords: *Automatic speech recognition, text summarization, sentence extraction, tokenization, parser, scoring, TF-IDF*

1. INTRODUCTION

Data is generated in both textual and audio formats. Webinars, technical talks, novels, e-books and journals are being published every day. A better method for efficiently and swiftly retrieving crucial information is therefore needed. Text summarization locates the crucial and meaningful information in a collection of documents and condenses it into a shorter version while preserving the overall meaning. It decreases the time required to read an entire page and the amount of space required to store big amounts of data. In single document text-summarization, a single document serves as the input, and information for the summary is taken directly from it. In the case of multiple documents, many documents are used as an input to generate the summary.

1.1**History of Audio Transcription**

Manual transcription has been in some form for few thousands of years. It has been largely benefited from artificial intelligence. Written representations or transcriptions of audio are helpful to readers without listening to the recording again. Transcriptions are required to preserve records, disseminate information, and to improve the accessibility. Advancements of Artificial Intelligence (AI) have led to Automated Speech Recognition (ASR) to aid with transcriptions. ASR systems can quickly transform human voice to text, and is one of the evolving research areas. The manual technique of audio transcription involves a human taking notes about words or occurrences in a meeting or event at the earliest possible. A person can listen to an audio recording of the event remotely and transcribe the same. In manual transcription, original notes can be referred and their transcription can be refined further. This method can provide high accuracy level, but it can be time-consuming and tiresome for the person involved. AI-powered transcription decreases the time for this process by handling the first transcription in real-time. ASR provides an easy mechanism to enter words into a document without delaying. Many people seek out its use because of its rapidity when compared to manual printing.

1.2. Text Summarization using Deep Learning

Feature learning methodologies are used to handle many DL challenges including word embedding. In Deep Learning, these representations are commonly referred to as word embedding. Word representations are downloaded and utilize for any application. However, each word in conventional neural embedding has a fixed representation regardless of context, which is a drawback. For example, the term *season*, may be used as both noun and verb depending on the situation and context. Several applications have recently replaced traditional word embedding with new pre-trained language models, resulting in considerable gains across several DL workloads. In general, there exist two text summarization approaches, namely, abstractive and extractive. The extractive strategy, as mentioned in [19] extracts and combines the most important sections of a text to create a summary, whereas, the abstractive approach generates a summary.

2. RELATED WORK

This section discusses techniques proposed in the Literature for text summarization. Extractive summarization involves the process of selecting the most important sentence. It is possible to choose the crucial paragraph by using the statistical and linguistic characteristics of paragraphs. Abstractive summarization captures the prime idea and meaning of a certain piece of writing or material. The novel idea from the text is generated by doing a linguistic analysis. The output will consist of the most recent condensed text that highlights important details from the document. According to previously proposed criteria such phrase frequency, word [8], significant phrases [2]. Research on extractive summarization is conducted in [8] where the essential sentences are retrieved by measuring word and phrase frequency, which provides a good indicator of their relevance. In [1], one crucial statement from the summary is extracted using two characteristics namely position and word frequency. Query-oriented summaries construct a query set based on the user's preferences. Background summaries are brief and attention-grabbing because they presume the reader has little prior information. Just-the-news summary offers readers with up-to-date news based on their interests. Single document summary uses sentences from the document, whereas multi document summary creates a summary by combining sentences from many papers. Topic-oriented summarization is based on user interest and information is retrieved from a given document related to the topic [1-2].

3.PROPOSED METHODOLOGY FOR AUDIO TRANSCRIPTION AND TEXT SUMMARIZATION SYSTEM

The proposed approach uses extractive method for summarizing the text. The extractive method chooses the text's most important sentences and outputs those to the user. The system is separated into two major components: transcription and text summarization. The design of the proposed Audio Transcription and Text Summarization System is illustrated in Figure 1. The user will be able to offer feedback in two ways. Through the front end, the user may upload WA3 format audio files or text files. First, the audio

ISSN 2394 - 9554

is divided into numerous distinct files based on frequency differences. The split audio files are stored locally and subsequently retrieved and transcribed. The audio files are fed into speech recognition module which will assist in recognising the speech and converting to text format. The transcriber module employed is a Google speech recognition module that recognises user commands in Google Assistant. The biggest drawback is that the audio recognition module can only identify one language in a file. The main summarizer module receives the transcribed audio and summarises the text depending on compression ratio and text line count.

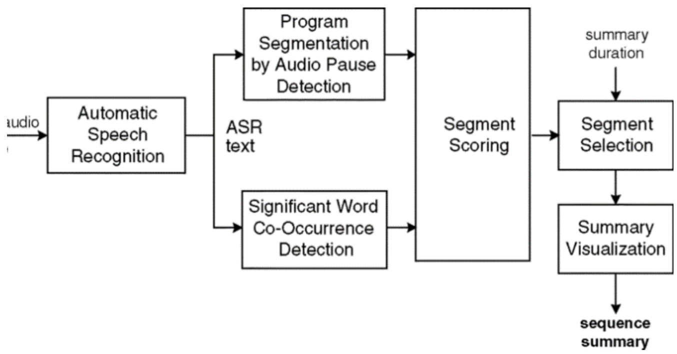

Fig. 1. Block Diagram of the proposed Audio Transcription and Text Summarization System.

The proposed approach employs both predefined and self-made structures to summarize the content. NLTK modules such as Tokenizer and Lancaster Stemmer would be among the preset modules. The self-made structures employ two methodologies, namely, Cost Calculation and TF-IDF. The summarizer module is divided into two key sections. The text is handled differently by each of the aforementioned submodules, and various calculations are used to determine the weights/scores for each phrase and word.

The scores from these several modules are combined to discover the best phrases that meet the aforementioned requirements. The webpage then displays the summary content. Text summarization is important in many day-to-day jobs since it enhances efficiency in grasping issues and substantially decreases time. Pre-processing of the input data is important in the technique outlined since damaged audio or audio captured in a noisy environment will not produce good results. So, breaking the audio into many parts would assist to alleviate the majority of the difficulties. Splitting is based on changes in frequency and the amount of silence in the audio. As the audio is split into smaller chunks, different audio files have different information inside them; therefore, if any faulty audio is present in it, will be ignored during the speech to text recognition process. This manner, the most critical and important information is kept while the sections are disregarded. After converting the audio to text, we employ NLP methods to pre-process the text before sending it into the summarizer module. Since summary primarily focuses on how frequently a term appears, the biggest issue would be handling of *is, was, but, if,* etc.

The probability of the stop words occurring are more compared to the most important words which are nouns or actions that someone performs. To handle this, a Python function is used to remove the stop words from the documents and will not consider them for scoring. By this way, only the frequency of the important words is considered for summarizing. After this, the rest of the words are tokenized. Tokenization would help uniquely identify a word and vectorise it based on the importance in the document. Tokenization would sometimes help in splitting the words and vectorise it. Vectorization is important in the process of summarization to consider the weightage of the words. Only after all the pre-

International Journal of Research in Science and Technology
Volume 12, Issue 01: January - March 2025
ISSN 2394 - 9554

processing is done, the document is fed into the summarization model. The summarization model then uses the pre-defined logic to summarize the text and display it on the front-end. The TF-IDF technique is widely preferred as it requires less time to process and produces comparatively better results.

4.AUDIO TRANSCRIPTION USING DEEP LEARNING

This section discusses about how the audio is transcribed into text by using various techniques. Transcriptions are written versions of audio information that help the reader to consolidate without listening to the recording multiple times. Transcriptions are required for the preservation of records, the dissemination of information, and the improvement of accessibility. As artificial intelligence (AI) advances, people are increasingly depending on a method known as automated speech recognition (ASR) to aid with transcriptions. ASR systems can quickly transform human voice to text, and their industry is already growing.

4.1 Audio Transcription

Retention and resource selection is the primary challenge. Finding transcribers and specialists may be difficult depending on the timeframes and the language of transcription. Other elements are some of the challenges while transferring the audio content. The speaker's vocal quality such as the audio's nature and content, the audio quality, and various dialects and accents. Because unscripted conversations make up the majority of the files for transcription, the quality may vary and only a small portion of the content may be useful.

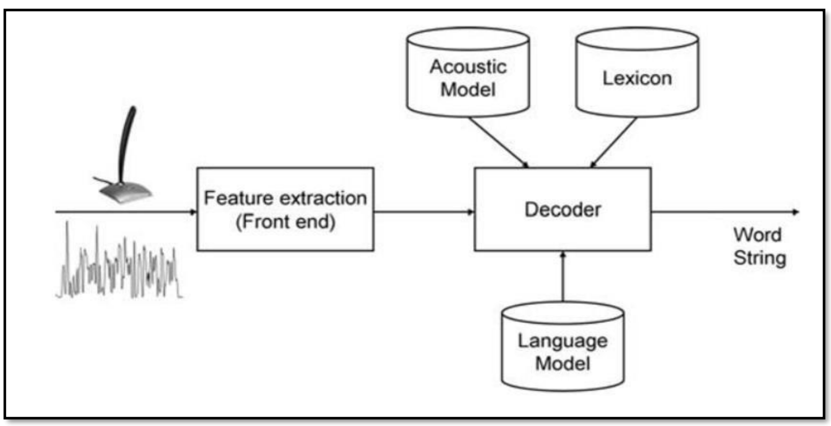

Fig. 2. A Typical Automatic Speech Recognition System

4.2 Automatic Speech Recognition

Automatic Speech Recognition (ASR) application converts spoken words into text and can identify group or individual dialogue. The words it hears are then turned into a wav file by the gadget. To remove background noise and balance loudness, the wave file is cleaned. This waveform is then disassembled and successively evaluated. The automated voice recognition software examines these sequences and evaluates statistical likelihood in order to identify full words and eventually complete phrases. There are two techniques to training ASR systems, whether they are NLP or guided conversation systems.

ISSN 2394 - 9554

Human Tuning: This approach of ASR training is simple and involves human programmers searching through specific ASR software interface's conversation record for frequently used words that the software interface required without previously being part of its pre-programmed grammar.

Active Learning: Applications developed with active learning feature detect user's specialized speech habits and communicate with them effectively. As a result, if an user continually rejects autocorrect for a certain term, the NLP application continuously learns to recognize that person's unique signature of the word as the correct form. A statistical method is used by the vast majority of current speech recognizers. Let A represent a string of words, W = w1, w2,..., wj,..., wn, and W represent a series of acoustic characteristics, A = ai, a2,...a3,..., ai, where each *ai* is a feature vector as mentioned in Equation (1). The primary purpose of speech recognition is to find the word string W′ that maximizes the likelihood of p(W|A).

$$p(\mathbf{W} \mid \mathbf{A}) = \frac{p(\mathbf{W})\, p(\mathbf{A} \mid \mathbf{W})}{p(\mathbf{A})} \quad \ldots \ (1)$$

Because p(A) is a constant that is independent of W, it may be ignored when looking for W′ as mentioned in Equation (2)

$$\mathbf{W}' = \arg\max_{\mathbf{W}} p(\mathbf{W})\, p(\mathbf{A} \mid \mathbf{W}) \quad \ldots \ (2)$$

The probability of a speech characteristic is given by *p(A|W).* While *W* is spoken, it uses the lexicon to detect word pronunciations and the acoustic model to model sound units i.e., phonemes based on speech attributes. P(W) is the priori probability of *W* and may be divided using Equation (3).

The approximation in the formula is a realistic technique to decrease the language model's complexity. An n-gram is a linguistic model in which the likelihood of uttering the word Wj is determined only by the previous n-1 syllables.3-gram is a frequent choice in many real-world speech recognition systems.

$$p(\mathbf{W}) = \prod_{j=1}^{J} p(w_j \mid w_1, \ldots w_{j-1}) \approx \prod_{j=1}^{J} p(w_j \mid w_{j-n+1}, \ldots w_{j-1}) \quad \ldots \ (3)$$

The ASR model consists of two main parts, the front end and thedecoder, as shown in Figure 2. The waveform of the voice is extracted via the front-end block's spectrum representation. The decoder block makes use of the acoustic model, lexicon, and the language model to find word sequences that match the incoming acoustic data.

4.3 Features of Automatic Speech Recognition

Voice recognition software is used because of its quick response. On the other side, typing could cause the process of communication to lag. Voice User Interface (VUI) is constantly evolving and has advanced significantly from earlier contact center technologies. Voice recognition and speech activation are primarily used in offices, where it may support and aid task-management. As technology progresses, asking a query or obtaining data for every specific instance will become more prevalent and can aid communication between persons who speak various languages by translating what is said in a foreign language into the recipient's native language, allowing them to comprehend and avoiding any language barriers in their regular business activities.

ISSN 2394 - 9554

4.4 *Natural Language Processing*

The field of artificial intelligence known as NLP focuses on the processing of natural language (AI). Across various fields and sectors, this technology for personal assistants is in demand. This technology examines a user's voice, deconstructs it for accurate comprehension, and then processes it.

Speech Recognition: Artificial language must first be created by the computer from natural language. This is accomplished through speech recognition, also known asspeech-to-text. This is the initial stage in the NLU procedure. Hidden Markov Models (HMM) are currently used in the most of the voice recognition systems. These statistical models translate your speech into text by figuring out what you said using mathematical computations. In order to achieve this, HMMs listen to you talk, break it down into small units (about 10–20 milliseconds), then compare it to previously recorded speech to identify which phoneme you said in each unit. A phoneme is the smallest unit of speech. The algorithm then evaluates the phoneme sequence and employs analytical statistics to the much more probable phrases and words should be chosen.

Natural Language Generation (NLG) is a lot simpler to use. Using text-to-speech technology, NLG turns a computer's artificial language into text that can subsequently be read aloud. Which data sets should be translated to text initially is decided by the NLP system. If the computer is asked a question about the weather, it most certainly will look up the answer online before providing the current temperature, wind speed, and humidity. Then it is structured according to how it will be said. Similar to NLU, but the opposite. The NLG system uses a vocabulary and a set of grammar rules to create entire sentences.

4.5 *Stemming*

The process of stemming entails developing morphological variations of a root or base word. The terms stemming algorithms and stemmers are used to describe stemmingprograms. The words "chocolates," "chocolatey," and "Choco" are shortened to the root word "chocolate," and "retrieval," "retrieved," and "retrieves" are shortened to the stem "retrieve." Stemming is an essential part naturallanguage processing. The stemmer receives tokenized words. Tokenization is the process of dissecting a document or a text into individual words. Figure 3 gives an example of how the stemming works.

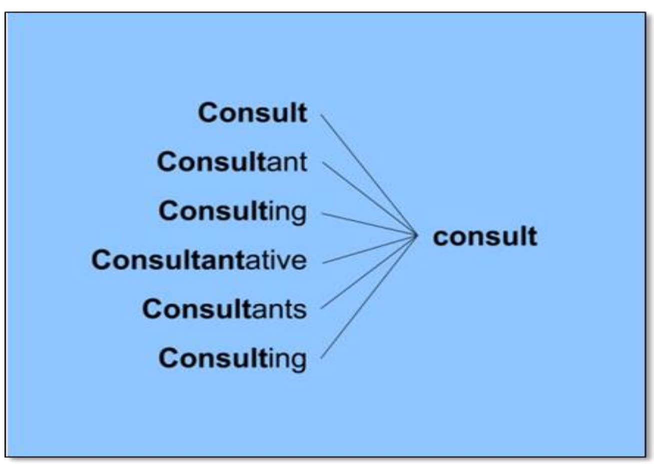

Fig. 3. Example for Stemming

ISSN 2394 - 9554

5.TEXT SUMMARIZATION USING DEEP LEARNING

The Previous Section presents the Audio Transcription isdone using the deep learning whereas this section discusses the text summarization using deep learning Techniques. The technique of breaking down lengthy publications into digestible paragraphs or phrases is known as text summarization. The proposed methodology extracts vital information as summary while also preserving the context. This reduces the time required to grasp lengthy items such as research papers while not compromising vital information.

The process of text summary conducted by robots or Artificial Intelligence algorithms is called as automatic text summarizing. However, there are certain disadvantages to using computerized text summary. The first challenge is selecting which information from the main source is relevant. The final summary must subsequently be presented in a reader-friendly way by the summarizer. By solving these challenges, automatic text summarizers aspire to increase subject coverage and readability.

There are two forms of text summarization based on outcome, as mentioned in [20] namely, extraction-based and abstraction-based. Further, text summarization can be classified into three types based on the context used: Domain-Specific Information, Query-driven, Generic methods of summarization. A few evaluation criteria are used to ensure that text summarizers attain the primary objectives of Optimal topic coverage and readability. One of the criteria is maintaining the most important component called salience. A summarizer must capture the most important information from the original material. The final summary must be precisely the appropriate length. A framework that is user-friendly is required. The sentences must be logical and coherent with balanced pronouns.

Developing a trustworthy method for text summarizers to comprehend what is vital and relevant is still a challenge, even with highly developed tools for writing and analysing summaries. Although word correlations can be found using vector representation and similarity matrices, there is still no reliable way to determine the most important phrases. Another challenge in text summarization is the complexity of human language and how people express themselves, particularly in writing. Language is made up of long sentences with adjectives and adverbs to describe something, relative phrases, and appositions. Though adjectives, adverbs and appositions are helpful in gaining the insights, they do not aid in determining the main point of data to be presented in the summary.

Term Frequency-Inverse Document Frequency (TF-IDF): TF-IDF is a metric that indicates the significance of a word in a document. The TF-IDF value rises in proportion to how frequently a word appears in the text. The sentence-frequency reveals how frequently the phrase is used in the text. After these sentence vectors are compared to the question, the summarization is affected by the inquiry since only the sentences with the highest scores are included.

TF/IDF score is calculated using Equation (4).

$$TF/IDF(w) = DN\left(\frac{\log(1+tf)}{\log(df)}\right)\Big) \quad \ldots \quad (4)$$

Cluster based method: In this technique, a collection of triplets—subjects, verbs, and objects—associated with each sentence apprehends and communicates in natural language the semantic core of a certain document. These triplets are grouped together to generate the summary.

ISSN 2394 - 9554

Graph theoretic approach: In this strategy, each sentence has its own node. If two phrases have similarity or share certain common terms to form an edge, then they are sufficiently similar. This graphic results in two things: The graph's divisions, or the sub-graphs that are not connected to other sub-graphs, produce distinct themes. The identification of important sentences in the text is the second outcome.

Automatic text summarization based on fuzzy logic: Each text attribute such as sentence length, resemblance to key word etc. is fed into a fuzzy system. In its knowledge base, the system keeps all the rules necessary for summarising. Following that, each sentence in the output is assigned a value between zero and one based on knowledge base rules and sentence features. The value obtained in the output determines the distinction of the sentence in the final summary.

Each and every sentence in the text is assigned a weight/score using the mentioned techniques. The finest sentences are chosen based on the summaries of the scores to create the summary that is subsequently published on the webpage. The summarization of information is useful in many day-to-day tasks that improve the efficiency in understanding topics and reduce the time drastically. As the summarization is mostly based on frequency of the word occurring, the most important challenge would be to remove the stop words such as *is, was, but, if,* etc. The probability of the stop words occurring are more compared to the most important words such as nouns or verbs. Therefore, the stop words are eliminated from the documents and will not be sued for generating the summary. The TF-IDF technique is popularly used to generate the summary as it takes less time to process and produces comparatively better results.

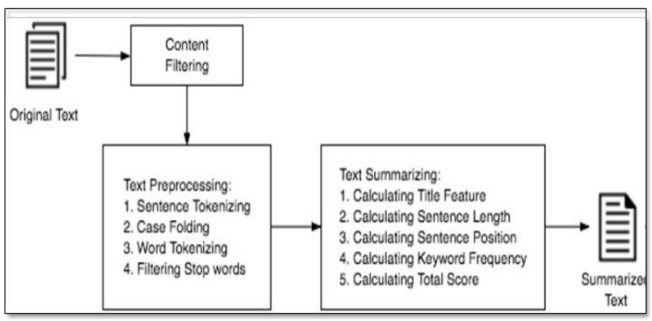

Fig. 4. Steps Involved in the Proposed Text Summarization Technique.

This section discusses about how the text is summarized using TDF IF and Natural Language Processing to convert the whole novel to summarized readable text. The next chapter analyses the data collection and its processing.

6.DATASET COLLECTION AND PREPROCESSING

The input datasets can be in textual format or audio format. The majority of available text-summary-datasets consist of short-form source texts with significant layout and stylistic biases and lack of long-range causal or temporal linkages. The BookSum [18] dataset consists of extremely abstractive, manually drafted summaries at three increasingly demanding degrees of granularity: paragraph, chapter, and book levels. Additionally, it draws on literary sources like plays, novels, and short tales. Th scope and structure of the dataset under consideration present a special set of challenges for summarization algorithms, including handling very long texts, complicated discourse structures, and non-trivial causal and temporal links. The collection of datasets for audio to test the transcription module was also done through gathering data from seminars and websites. The audio gathered includes clean audios, noisy audios, audios with frequency being tampered and audios with intentional breaks in between. Pre-processing of data is important for both audio transcription and text summarization. The preprocessing of audio

ISSN 2394 - 9554

transcription is done to improve the accuracy of the transcription. The audio file is split into multiple mini-audio files if there exists a gap of 1000ms between two different words to improve the efficiency. The audio files are split and temporarily stored in the local host and the files are transcribed one by one using the transcription module. The pre-processing for text summarization is done separately in the summarization phase. NLP techniques is used to remove the stop words from the sentences thereby only the important words are considered for scoring and summarizing.

7.METHODOLOGY

The major part of the system is the summarizer module which is made up of two sub-modules: TF-IDF and Cost Calculation modules. Term frequency–inverse document frequency (TF-IDF) represents the importance of a word in a document. The TF-IDF value of a term will be higher if the term appears more frequently in the document. The sentence-frequency is the number of times a term appears in a sentence. The summary is constructed using the sentences that received the highest score. The above-mentioned terms are computed using the Equations (5), (6).

$$Tf(t, d) = count\ of\ terms\ in\ d\ /\ total\ number\ of\ words \quad \ldots(5)$$

$$idf(t) = \log(N/(df + 1)) \quad \ldots(6)$$

In order to determine the cost, two sentences are connected by an edge if they share words. The cost calculation method's second outcome is the identification of the document's significant sentences. Figure 3 shows the logic and idea behind how the cost calculation method works. The most significant assertions in the partition have the largest number of edges linking them, and has high inclusion priority. The average cost is calculated and the edges below the average are removed. The rest of the words are assigned score based on the average. Finally, the scores are normalized and aggregated together to score each and every sentence.

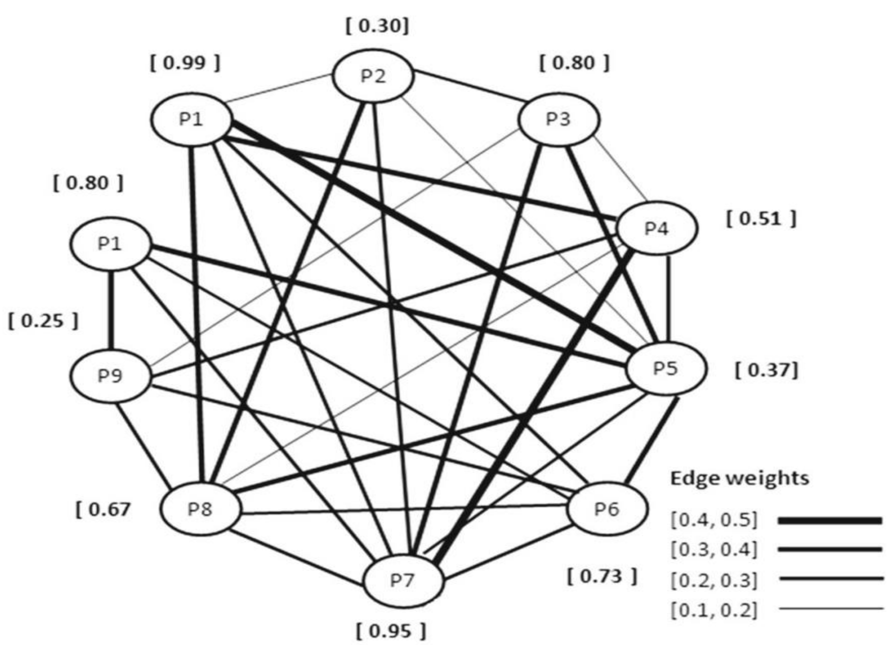

Fig. 5. Graph Model with Words as Vertices and Common Phrases as Edges.

 ISSN 2394 - 9554

The steps involved in generating the summary is outlined using the following algorithm

Procedure Text Summarizer ( )

{

Input: audio in the specified format

Output: extracted short summary

1. Convert the audio to text.

2. Pre-process the text and choose the important words.

3. Tokenize the processed words and provide it as an input to the summarization model.

4. Each token has a part of speech associated with it. Individual tokens are given weights using TF-IDF score.

5. Frequency of each and every word in the document is divided by the maximum frequency to determine the weight frequency of the tokens.

6. The ranks for individual terms are calculated and the summarizer will ultimately extract the high weighted frequency sentences based on the user's input and summarize the text.

}

8.EXPERIMENTAL RESULTS

The proposed methodology is tested with a 20-minute TEDx audio on 'Become the Person you can't imagine' by Norman Bascal. Figure 6. shows the sample text that has been recognized from the audio input.

```
audio-chunks\chunk1.wav : Next few minutes i don't want to change the blueprint.
audio-chunks\chunk2.wav : Have your life your career or even you tomorrow.
audio-chunks\chunk3.wav : Price duke 125 ke bad.
audio-chunks\chunk4.wav : What i've learnt has changed my tomorrows my career and my life.
audio-chunks\chunk5.wav : Yesterday i want to introduce e27 to share with me today it is unko harry.
audio-chunks\chunk6.wav : You can see him all the pictures.
audio-chunks\chunk7.wav : It was originally a lot creare yes is p**** moustache with a white lab coat with stethoscope brands n
eck and always has this and which cigarette smell.
audio-chunks\chunk8.wav : Ishwar with distinction mcnairy navy and watch two and one or two become one materials must be loved
pediatricians.
audio-chunks\chunk9.wav : House roof patients thermacol harry in fact many i'm still there.
audio-chunks\chunk10.wav : For 19 years old just returned from your authentication with great han pusi haircut and long beard r
eady to begin my major in biology mcgill university.
audio-chunks\chunk11.wav : Adwords.
audio-chunks\chunk12.wav : Hona chahie.
audio-chunks\chunk13.wav : B y gauri shankar harry free advice he being a doctor now.
audio-chunks\chunk14.wav : Nobody tells me.
audio-chunks\chunk15.wav : Is not what i expected.
audio-chunks\chunk16.wav : Normal.
audio-chunks\chunk17.wav : Mrs.
audio-chunks\chunk18.wav : Think of your career as a weather.
audio-chunks\chunk19.wav : Aap jois.
audio-chunks\chunk20.wav : You can watch about in water at paro upstream ago you think you want.
audio-chunks\chunk21.wav : Only the find you right there exhausted and it is virtue.
audio-chunks\chunk22.wav : You can steer your belt and stream.
audio-chunks\chunk23.wav : Find out the bans entrance which provide you opportunities in never could have imagined.
audio-chunks\chunk24.wav : Kachora idea most incredible.
audio-chunks\chunk25.wav : One more thing known.
audio-chunks\chunk26.wav : Every seat you make sure you come together.
audio-chunks\chunk27.wav : Well.
audio-chunks\chunk28.wav : Stationery supplies so what i do with that i did but also to university students with i gathered tog
ether with my buddies.
audio-chunks\chunk29.wav : Environment.
audio-chunks\chunk30.wav : Kisi rocky horror picture show.
audio-chunks\chunk31.wav : Teen x.
audio-chunks\chunk32.wav : Now i never discussed ago harry's advice with anyone for a 4 years.
```

Fig. 6. Transcription of Digital Audio-1

 ISSN 2394 - 9554

The accuracy of the text is quite low since the audio was recorded in a noisy room. Although the distortion of audio is high, the module still manages to extract most of the information. The first chapter from the novel 'The Jacket' by Jack London is fed as a text-format input to the model. Figure 6 shows the first chapter of the novel for context.

The summarizer is then tested based on the amount of information it has to retain. The number of inputs is can range between 1-100. The summarizer shows distinct results based on the input provided. The summarized results are tested using the Rogue evaluator. Recall-Oriented Understudy for Gisting Evaluation (ROUGE) is primarily a set of evaluation tools for machine translation and automatic text summarization. Sample summary generated is shown in Figure 7. ROUGE compares an automatically generated summary against a group of manually compiled reference summaries. The summarization precision is determined by precision, recall, and f-measure metrics. Recall indicates that the system summary has caught all of the terms in the reference summary. Few test cases are differentiated based on the amount of information to be retained.

```
Summary: We new-born infants, without experience, were
born with fear, with memory of fear; and memory is experience. Yes, I,
whose lips had never lisped the word
"king," remembered that I had once been the son of a king. More—I
remembered that once I had been a slave and a son of a slave, and worn
an iron collar round my neck. Silly, isn't it? I was caught red-handed.
The court records show that I did; and, for once, I agree
with the court records. Men who endure it, call it living death. I can
pick the high percentage butter-fat cow with my eye and let the Babcock
Tester prove
the wisdom of my eye. Many a farmer, riding in his motor-car
to-day, knows who made possible that motor-car. And farm management!
```

Fig. 7: 10% of Information is Retained as Sumamry

Table 1. Recall vs Information retained

Information Retained	Recall
30%	0.296538462
50%	0.530769231
10%	0.096153846

Table 2: Precision and F1-measure

	Information Retained	Precision	F1-measure
Rouge1	10%	1.0	0.175439
		Precision	F1-measure
Rouge1	30%	1.0	0.414139
		Precision	F1-measure
Rouge1	50%	1.0	0.693467

The Table (1) shows the Recall Vs Information Retained, which is about 98% accurate. The information retained is always dependent on the user input. To calculate the precision and F-measure, instead of comparing the summary that was produced automatically and a list of reference summaries, it is compared against the original text. The Rouge-1 metric is used and the Table (2) shows the Precision and the F1-measure for different test cases. The F1 score considers both accuracy and recall and strikes a compromise between the two. Thus, high F1 scores will be marked for accurate predictions. The F1 score is calculated using Equation (5).

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \qquad \ldots \quad (5)$$

Information to be retained decides the precision of the summary can be inferred from the table (2). The information retaining parameter is directly proportionate to the precision and the recall value. These parameters have cleary indicated that even though the model is an automatic text summarization model, the accuracy of the moel will be determined by the user. This is because the user will be able to provide the amount of information to be retained as an input and since the information retaining parameter is directly proportionate to the precision and the recall value, the performance and accuracy of the model is dependent on the user's input.

9.CONCLUSION AND FUTURE WORKS

The complexity of the automatic text summarization is broken down by dividing the process into numerous smaller tasks. Every subtask has the potential to provide high calibre summaries. The key to extractive text summarization is identifying the critical paragraphs. In this study, we suggested extractive-based text summarization utilising two separate models combined with their combined scores and ranking. Based on the ranking provided, the text is summarized and displayed to the user. When compared to more conventional methods, the suggested model is more accurate because it incorporates user input to do so. To further enhance, an Android mobile application can be created, and this model can be integrated with it to improve the project. A camera can be utilized in the mobile application to recognize the text and describe it. Here, summarization may also be extended to other regional languages.

10. REFERENCES

[1] S. Deerwester, S. T. Dumais, T. K. Landauer, G. W. Furnas and R. A. Harshman, “Indexing by Latent Semantic Analysis”, Journal of the Society for Information Science, 41(6), pp. 391-407, 1990.

[2] H.P. Edmundson, “New Methods in Automatic Extracting”, in Advances in Automatic Text Summarization, an edited book, MIT Press, pp. 23-34, 1999.

[3] J. L. Fagan, "Experiments in Automatic Phrase Indexing for Document Retrieval: A Comparison of Syntactic and Non-syntactic Methods", Ph.D. Thesis, Department of Computer Science, Cornell University, 1987.

[4] J. Goldstein, V. Mittal, J. Carbonell, and M. Kantrowitz, “Multi-Document Summarization by Sentence Extraction”, NAACL-ANLP Workshop on Automatic Summarization, 10.3115/1117575.1117580, 2002.

[5] E. Hovy and C. Lin, “Automated Text Summarization in SUMMARIST”, in “Advances in Automatic Text Summarization”, MIT Press, pp. 81-94, 1999.

[6] F. Kelledy and A.F. Smeaton, “Automatic Phrase Recognition and Extraction from Text”, 19th Annual BCS IRSG Colloquium on IR, UK. 8th-9th, 1997.

[7] Liu, Shuhua, “Enhancing E-business-intelligence-service: A topic-guided text summarization framework. Proceedings” - Seventh IEEE International Conference on E-Commerce Technology, CEC 2005. 2005. 493- 496. 10.1109/ICECT.2005.45, 2005.

[8] Luhn, H.P., “The Automatic Creation of Literature Abstracts” - IBM Journal of Research and Development, 2, pp. 159-165, 1958.

[9] I. Mani and M. T. Maybury (Eds.), Advances in Automatic Text Summarization, MIT Edited Book, 1999.

[10] C. D. Paice, “Constructing Literature Abstracts by Computer: Techniques and Prospects”, Information Processing and Management, 26 (1), pp.171-186, 1990.

[11] D. Radev, T. Allison, S. Blair-Goldensohn, J. Blitzer, A. Celebi, E. Drabek, W. Lam, D. Liu, H. Qi, H. Saggion, S. Teufel, M. Topper and A. Winkel, “The MEAD Multi-document Summarizer”, 2003.

[12] G. Salton, A. Wong, and C. S. Yang, “A Vector Space Model for Automatic Indexing”. Communications of the ACM, pp. 613-620, Volume 18, Number 11, 1975.

[13] G. Salton, “Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer Reading”, Addison-Wesley, pp. 351 – 378, 1989.

[14] G. Salton, J. A. C. Buckley, A. Singhal, “Automatic Analysis, theme generation and Summarization of Machine Readable Texts”, 264, pp. 1421-1426, 1994.

[15] G. Salton, A. Singhal, A. Mitra and C. Buckley, “Automatic Text Structuring and Summarization”, Information Processing and Management, Vol. 33, No.2, pp. 193-207, 1997.

[16] P. Sheridan and A. F. Smeaton, “The Application of Morpho-Syntactic Language Processing to Effective Phrase Matching”, Information Processing and Management 28(3): 349-370, 1992.

[17] K. Spärck-Jones, “Automatic Summarizing: Factors and Directions”, Advances in Automatic Text Summarization, MIT Edited Book; 1999.

ISSN 2394 - 9554

[18] W. Kryściński, N. Rajani, D. Agarwal, C. Xiong, D. Radev, “BookSum: A Collection of Datasets for Long-form Narrative Summarization”, 10.48550/ARXIV.2105.08209, 2021.

[19] B. Hachey, C. Grover, “Extractive summarization of legal texts”, Journal of Artificial Intelligence and Law, 14, pp. 305–345, 2006.

[20] A. Nenkova, K. Mckeown, “A Survey of Text Summarization Techniques”, 10.1007/978-1-4614-3223-4_3, 2012.