



## Automated Classroom Lecture Note Generation Using Natural Language Processing and Image Processing Techniques

T.C. Sandanayake<sup>1</sup>, A.M. Bandara<sup>2</sup>

<sup>1</sup> Faculty of Information Technology, University of Moratuwa, Katubedda, Sri Lanka, thanujas@uom.lk

<sup>2</sup> Faculty of Information Technology, University of Moratuwa, Katubedda, Sri Lanka, anusha13@itfac.mrt.ac.lk

### ABSTRACT

Teachers use different teaching methods to deliver the course content to the learners. Smart learning environments helps to reform of the ways of teaching and learning through advancing current methods. These environments help learners to enhance their knowledge and skills by engaging in different learning activities. These different methods can be use of power point, white board drawings, and online learning platforms, audio and video inputs. Teacher's main aim is to maximize the learner engagement in developing knowledge and skills. The teachers' expectation is that the learner's to gather knowledge though the variety of lesson materials. This paper describes a research approach on how the learning materials, smart white board input and the lectures' voice input can be combined in generation of compiled lecture note. In developing the solution it has used the voice input, white board diagrams and other earning materials used in the classroom, materials given in the Learning Management System and any internet materials. The study has adapted Natural Language Processing and Image Processing techniques. Main components of the research were text segmentation, text summarization, topic modeling, and diagram recognition. Outcome of the research can be used by the teachers and learners in day today work.

**Key words:** Text Segmentation, Image Segmentation, Topic Modeling

### 1. INTRODUCTION

Digital learning is any type of learning method that is facilitated by technology or by instructional practice that makes effective use of technology [1]. At present many systems are used to make learning process easier. Many learners face difficulties of unable to collate the lecture notes properly due to many reasons. In the academic environment teachers uses many different types of methods to deliver the lesson content to the learners. Even in the regular face to face classroom environments, a teacher uses variety of methods to deliver lessons to the learners. Some of those teaching and learning methods are, use of PowerPoint slides, flip charts, smart boards, video and audio content. At the same time

teacher uses to explain the things which are not in any of those materials during the conduct of the class. This research study has developed based on real-time input extracted during the

class room conduct. The study has used to create a lecture note by capturing lecturer's voice, text, diagrams on board and extra lesson learning materials. Visual learners and auditory learners are the beneficiaries of this research work.

The research further evaluates and extracts the data which made available on the smart white board and the voice input of the lesson explanation of the lecturer. Therefore the research has identified the research problem statement as "How to develop a complete lecture material based on different lesson inputs provide by the teachers?". Based on the research problem statement the research methodology was identified.

The main aim of the research study was to "design and develop a method of generating a complete classroom lecture note by evaluating the lesson materials and teacher voice input". The outcomes of the research study help the learners to solve above problem of creating and compiling a complete lecture note without delay. In the research process the researcher captured the lesson explanation voice of the lecturer, diagrams on the smart white board and other lesson learning materials in order to provide a compiled lecture note by processing them.

The research study has conduct based on the different technologies such as natural language processing, image processing. Speaker's or lecturer's voice is generating in natural languages and Speech recognition methods [2] were used to identify that natural language. Diagram identification, sentence segmentation, example extraction and text summarization, topic modeling are the main target areas in this application. The automated combined application was designed and developed as the outcome of the research in order to derive the research aim. The research study has been carried out with the Level 3 and Level 4 undergraduate learners and the academic staff of the Faculty of Information Technology of University of Moratuwa. Research study has been tested and validated with the learners and teachers upon the physical observation of the class room

activities. The research results will be helpful for the future researchers who will be using the video output of the lesson to derive the similar outcome.

## 2. LITRETURE REVIEW

The research study has mainly focused on the NLP and Image processing literature. Sentence segmentation is one of the key literature which has been reviewed under the study. Sentence segmentation process is used to segment the whole text file into separate paragraphs. This text file is the output of, voice to text converting process. Hearst(1997) has introduced a method for this process as sentence segmentation using TextTiling [3]. It is an algorithm for breaking documents into topically coherent multi-paragraph subparts. Semantic similarity [4] matching is the other method that researchers have introduced for the sentence segmentation process. Measuring semantic similarity of a given sentences in any scenario is closely related to the mechanism on the method of checking semantic similarity between words. This process assures that it derives the relationship between a word and the sentence through the meanings. Vector Space Models [5] are also being used in segmentation and every word is mapped in a vector that contains other words with a score that represents the connection between that word and the others.

Text Summarization and Example Extraction methods are also being used in similar research approaches. Hence text summarization process is used to create a summary of clustered paragraphs. There are two main types of text summarization methods, as extractive text summarization and abstractive text summarization [6, 7]. Extractive summarization [8] is used to extract important sentences or phrases from the original document and group them to produce a summary without changing the original text. Abstractive summarization performs summarization by understanding the original text with the help of linguistic methods to understand and examine the text [9, 10]. Example extraction is a kind of sentence extraction method that extracts the example sentences. Co-reference Resolution [11] is the most effective method is used to extract sentences.

The next key literature which connects the research is topic modeling. Topics are succinct representation of content in a document collection and hence are very effective when used as content identifiers in peer-to-peer systems and other large scale distributed content management systems [12]. Latent Dirichlet Allocation Model [13] is used for the topic modeling process. Multinomial Mixture Model [14] which is a probabilistic generative model used for documents, and embodies two assumptions about the generative process as, the documents are generated by a mixture model and there is a one-to-one correspondence between mixture components and clusters.

The research study has used image processing techniques for the diagram Recognition. The main purpose of this activity was to provide a meaning to drawings on the smart board by the teacher. A neural network approach for diagram interpretation is proposed Kembhavi et al [15] and here the problem of diagram interpretation, the challenging task of identifying the structure of a diagram and the semantics of its constituents and their relationships is studied. Further Kembhavi [15] has used interpretation and reasoning in the context of science diagrams, defined as the two tasks of Syntactic parsing and Semantic interpretation

## 3. RESEARCH ME OBJECTIVES

The research consists of four objectives which are as follows:

- Identify a method to segment the text document by converting the speech input to text.
- Identify a mechanism to generate topics of the lecture material.
- Develop a method to identify the diagrams in the lecture.
- Design a mechanism to match the summarized text data and diagrams.

## 4. RESEARCH METHODOLOGY

The research methodology consists of several sub modules. The sub modules of the methodology are Voice to text conversion, Text segmentation, topic modeling and diagram identification and note generation. The overall architecture of the research study is given in figure 1.

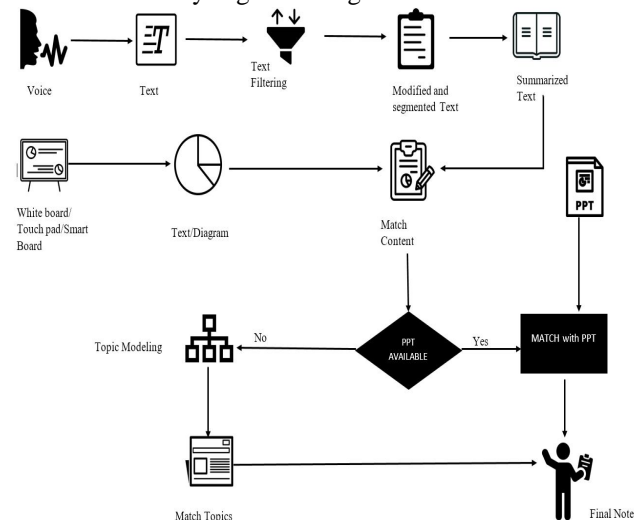


Figure 1: High Level Architecture of the System

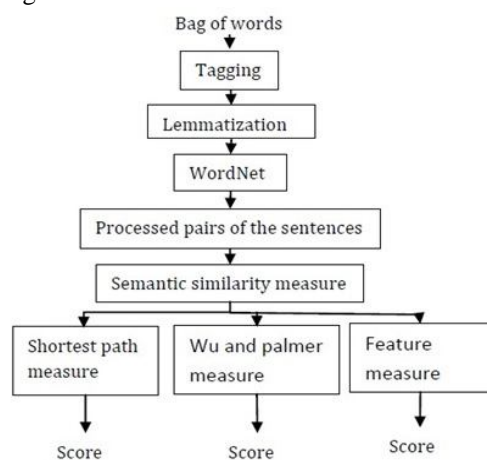
### 4.1 Voice to Text Conversion

The first module of the research was voice to text conversation. This voice input has been extracted during conduct of the lecture. There the lesson explanation has been captured in order to compile the lecture note along with other lesson materials such as power point slides and smart board

drawings. This voice should have to be converted into a text document. Google Cloud Speech - to - Text tool was used for this process. Here, lesson explanation audio file is included as the input of this process and converted text file is provided as the output. This text document is a some what pre-processed text document. That means, this tool identify the words correctly, sentence boundaries, and also add punctuation marks too.

## 4.2 Text Segmentation

In the second module, the text output from the previous step is separated into paragraphs. That text document contains all the details such as lesson introduction, objective explanation and other content details regarding the lecture. Most of the time it is not expected to get this output in a proper order to structure due to different reasons such as use of verity of learning materials and use of different teaching methods. Therefore during this process, was separated unstructured text document into paragraphs based on semantic similarity between sentences. Figure 2 shows an example for the classifying the text and sentences.



**Figure 2:** Flow of Semantic Similarity Computation

Hierarchical Algorithm for Sentence Level-Text Clustering - Clustering is the task of grouping a set of texts in such a way that texts in the same group (called a cluster) are more similar to each other. In this approach the main algorithm used for text clustering is hierarchical clustering. It has stated that the hierarchical clustering algorithms are designed either top-down or bottom up approaches. Here the research used hierarchical agglomerative clustering method. It works in a bottom-up manner. Here in each object is considered as a single-element cluster (leaf) at the beginning. At each step of the algorithm, the two sentences that are the most similar are combined into a new bigger cluster (nodes). In order to get the set of clusters that have similar meaning. A cluster represents a meaningful paragraph given as below.

Word Similarity Measure - Before merging sentences together it is important to consider about word similarity of

the sentences. Based on that result, next step is clustering the sentences. To measure similarity of words WordNet is used. WordNet is a lexical semantic dictionary which can be used to check semantic similarities. For this approach we use nouns of the sentences. This research approach the clusters initially will be based on Nouns such as 'manager', 'research', 'hierarchy', 'top level managers' and 'accounts'. After the similarity computation, hierarchical clustering is formed for the most similar nouns.

Sentence Similarity -Sentence similarity is the next step to be carries out in the research study upon completion of the word similarity. When finding out the similarity of the words then can determine how far sentences are similar to each sentences in given text document. Meaning of the sentence is reflected by the words in the sentences. Given two sentences, the measurement determines how similar the meanings of two sentences is. To measure the similarity of sentences this method dynamically forms the semantic vectors solely based on the compared sentences [5]. To find similar meaning words we used Jiang and Conrath's measure [16].Likewise using vector generating model, can derive all the vectors for sentences. Other important process is measure similarity between each and every vectors. For that we used cosine similarity measure.

Example Extraction - According to the content of the research study, there is a relevant place to display examples separately. When reading one paragraph, the examples which relevant to a particular paragraph are displaying along with the paragraph. This is a kind of sentences extraction method. Because it extracts the sentences that contain examples related to relevant lesson. When considering similar approaches the best method for this is co-reference resolution method. But here, it cannot be applied because; only few sentences are extracted from whole paragraph. Therefore, a new method was introduced using regular expressions [16] and word similarity. Regular expression is used to identify the words that used to describe examples; some are examples, such as, instead of, like and etc. First, according to given words, it catches the whole sentence starting from relevant word. Then it consider the sentence similarity and word similarity of the selected sentence and other sentences following it until getting probability of sentence and word similarity 0.

### A. Text Summarization

Text summarization process is used to summarize the segmented text data. Very important sentences was taken from the segmented paragraphs to proceed the lecture note. It is extractive text summarization. This is done by using word frequency and inverse document frequency methodology [8]. In the research, lesson learning materials are also considered as input data. These files contain abstractive information

about the lesson. Therefore, abstractive text summarization method was not taken as in this research approach. Here, segmented paragraphs are taken as the input data of this process. Then set of summarized sentences are given as the output of this process.

### B. Topic Modeling

This is the research module for topic modeling process. Above summarized paragraphs need a topic for clear identification of the lesson. Then, this module generates the relevant topics for each summarized paragraphs using new methodology. The clustered text is taken as the input data of this process. In this process, first step is to check the availability of lesson learning materials. If there are any files like pdf or ppt documents, all the topics are caught from related document. It is done by identifying key words of each paragraphs and each topics. LDA model used to identify these key words. Then, above paragraphs are matched with these captured topics. Cosine similarity measure is the other supportive method used in this process. When it does not contain lesson supportive documents, topics are generated using above technologies.

### C. Diagram Recognition

Here the diagrams drawn on the board are matched with the text paragraphs. A diagram may contain words, sketches. In this case it was mainly consider diagrams which have text content [17]. To extract words from the diagrams it has been followed two steps. First is the recognition of the text region in the image and the second is identifying the characters of those text regions. For identifying text region, an image processing approach is used. K-nearest neighbor classification is used to identify the characters. The mainly used concept is the connected components. The input image is put through binary thresholding as the first step. The resulting image is then subjected to morphological dilation. Connected component analysis is done on the result obtained from the previous step. Then the identified connected components are filtered according to predefined maximum and minimum values [17]. The successful candidates are cropped from the image and forwarded for character recognition. All this steps are done using python open cv.

Image Classification - K-nearest Neighbors Algorithm - k-nearest neighbors algorithm (k-NN) is a non-parametric method used for classification and regression Output from the text region extraction process is used as the input for this process. Initially an image processing is run to extract characters from the input image [18]. Then a pre trained model is used to identify characters. Prior to running the application it was place a training classification file. Images of handwritten characters are used for that training purpose .K-nearest neighbor is a classification algorithm used for supervised learning [18]. The research study was used the k-nearest neighbor implementation provided with open Cv.

## 5. RESEARCH RESULTS

Outcomes or results of each module could be able to achieve the objectives of the research. The sentence segmentation was the first step of the research outcome and it has succeeded about 72% of the outcome generated. When it gives a single text documents with contain different ideas, the application divides it into different segments according to meanings.

### 5.1 Text Summarization and Example Extraction

Results of this module give the achievement of the given objectives. Text summarization is done by using term frequency and inverse document frequency method. Word counts of the summary also generated according to frequency of the keywords in the text. Therefore, this method also achieves its goal related to the approach. Given Figure 3 and Figure 4 are the results of this module.

An ecosystem is a community made up of living organisms and nonliving components such as air, water, and mineral soil. Ecosystems can be studied in two different ways. Ecosystems can be thought of as interdependent collections of plants and animals, or as structured systems and communities governed by general rules. The living (biotic) and non-living (abiotic) components interact through nutrient cycles and energy flows. Ecosystems include interactions among organisms, and between organisms and their environment. Ecosystems can be of any size but each ecosystem has a specific, limited space. Some scientists view the entire planet as one ecosystem

**Figure 3:** Input Text of Text Summarization Process

eco system can be through of as interdependent collection of plants and animals, or as structured systems and communities governed by general rules. The living bodies and non living components interact through nutrient cycle and energy flows.

**Figure 4:** Results of Text Summarization Process

The next activity was the example extraction module. Here the When it considers example extraction method, it has achieved about 90% of effective results by extracting relevant examples from the given text. Figure 5 and Figure 6 gives the outcomes.

Evidence has accumulated indicating that various personality characteristics, such as locus of control and optimism, are related to how people cope with stress. For example, an optimistic orientation has been associated with increased problem-solving efforts, especially in controllable situations. Also, internal locus of control beliefs have been found to be associated with increased problem-focused coping or more adaptive coping. A major function of secondary appraisal is to determine what can be done about a stressful event, or whether it is controllable.

**Figure 5:** Input Text of Example Extraction Process

Optimistic orientation has been associated with the increased problem-solving efforts, especially in controllable situations. Also internal locus of control beliefs have been found to be associated with the increased problem focused copying or more adaptive copying.

**Figure 6:** Results of Example Extraction Process

## 5.2 Topic Modeling

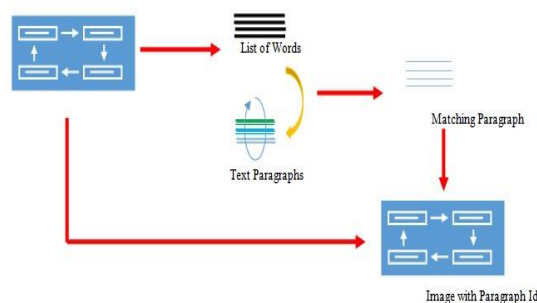
In the topic modeling process the whole text was summarized as paragraphs in order to derive the topic for the paragraph. Here generates topics for each summarized paragraphs with the help of two main methods. It is done by using two methods as, the situation which is having PDF document and which is not having PDF document. In this process, first consider the key words of the given paragraph using word frequency and then check the availability of a PDF document. According to the availability of PDF document each topic modeling methods are applied. The key assumption in this process is that the facts (key words) relevant to the topic, must be spoken in the content (speech).

The process of modeling the PDF files as follows. Assume that here one letter is representing a one word. First get the preprocessed word list and then select the word which is having highest value from the generated key words. Next is to select the topics from the topic list that are relevant to selected word. Then select the sentences from the paragraph that are relevant to selected word. The final step is to check the cosine similarity between above selected topics and sentences and get the topic which equal to the sentence with high cosine similarity as the most suitable topic.

The topic modeling method carried out for the non PDF files. The assumptions in this modules are; sentence that select as the topic having least value of length, the word which is having highest frequency in the paragraph has a high probability to be the topic, The sentence that take as the topic should be in the top level of the paragraph. The steps of the process are as follows. Get the probabilities of the words. Select the words which are having highest probability and select all sentences that contain selected words from the paragraphs. Filter the sentence according to the length. When there are dissimilar words in the sentence, assign "0" to the words. When there are words less than length, assign  $1/\text{length}$  value.

## 5.3 Diagram Recognition

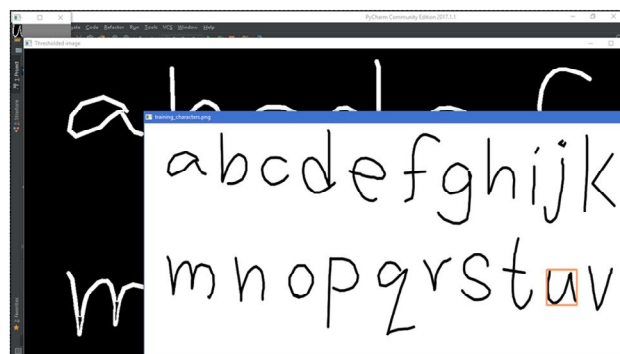
Diagram extraction is one of the key activities conducted in the research study. Here the research was identified the diagrams related to the lecture using image processing techniques. The lecturer is usually drawing the diagrams related to the subject content on the smart white board and the module recognizes the image using image processing techniques. This image identification was helpful for the learners to connect the topics and relevant lesson content with the identified image. In this process an android application was developed to use as the white board to draw diagrams. Diagram recognition module has been described in the Figure 7.



**Figure 7:** Diagram Recognition Process

The diagram recognition process has been described using two modules. The two modules are diagrams drawn on the board and text which are segmented into paragraphs. First, the text in the diagrams are extracted as words. Then for each image, a search step is run in order to find which paragraph contains the extracted words of the image. Accordingly the diagrams are matched with the paragraphs described by the text. To reduce the search scope for a given image, the study has used the time range in which a particular paragraph is spoken and the time in which the image is captured.

**Classifier Training:** Classifier is trained using a dataset which contains images of handwritten English alphabet (both simple and capital) and numbers from 0 to 1. Some of the images are collected by us and others are from various online resources. Corresponding input is given when the character is prompted. Figure 8 shows the process clearly.



**Figure 8:** Classifier Training

**Text Region Extraction:** The possible text regions are cropped out from the image in this phase. Those cropped images are then forwarded for character recognition. Figure 9 gives the image f text region extraction.

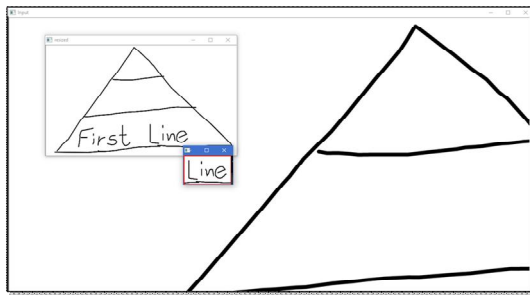


Figure 9: Text Region Extraction

Character Recognition: The possible characters are identified from the cropped images of the previous step. Characters from such an image are considered as a single word. These words are saved into a text file which has the same name as the input image. Figure 10 gives how the characters of the given images were recognized.

When diagram and paragraph matching, for key words in each paragraph in the text (speech to text output), a search mechanism is run in order to check whether that key word is available in any image (by using the text files created for each image in the previous character recognition phase). If it is available, then the relevant image link is attached to the paragraph. In order to reduce the search scope and increase the efficiency, only the images that are created within the same time range as of the paragraphs are considered.

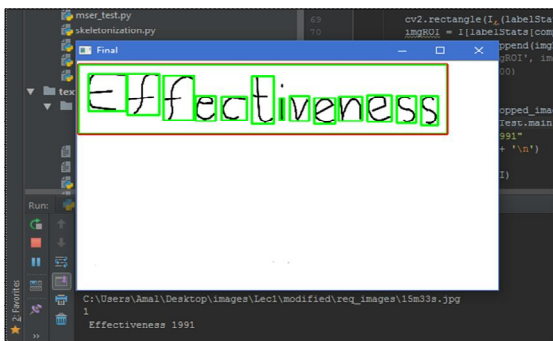


Figure 10: Character Recognition

#### 5.4 Application to generate final lecture note

The proposed research solution was developed as a software application in order to assist the learners. In this research application user can customize the note according to user's preference. Therefore it contains different options to add for smart note. That means, if user wants to take example or activities of the lesson separately he can select the given check box and crate the note. That is the way of use this research application. Following image presents the output of the smart note.

Figure 11 shows the output of the software application. When it selects one topic from the content, the relevant note for selected topic is displayed in smart note position. If user wants

to get examples separately above example position displays the examples related to selected topic. As well as activities of the given topic and relevant diagrams of the relevant topic are displayed in the related positions.

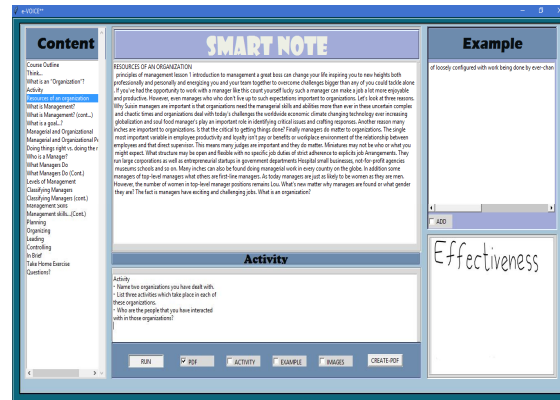


Figure 11: Smart Note Software Application

The Figure 11 gives the final outcome of the smart note research application. This software has been used in the purpose of validation of the model derived from the research. The framework has been used with the some of the course modules with the help of students and the lectures. As given in the screen shot of the software application the screen provide the access to the content as given in the left hand side of the screen appears as a 'table of content' of the lesson conducted by the teacher. This has been listed the topics, subtopic and activities .The learners are able to navigate through the topics and select the content what they want to appear on the lesson note. At the same time the activities done under each topic appears at the bottom of the screen with relating to the topic and subtopics. The images used in the lesson are appearing in the right hand side corner of the screen respectively. Finally the learners are having the freedom to use this research work to create their own lesson notes according to their requirements.

## 6. CONCLUSION

The learners face difficulties in collating all the lecture materials and what the teachers does in the classroom due to heavy workload. The research study was aiming on developing a method to generate the summarized lecture material based on the voice output of the teacher, white board diagram and any material (printed form). The main research methods were used a Natural Language Processing and Image processing technologies. The research was mainly limited to the classroom teaching environment but not via online. The main aim of the research study was to design and develop a method of generating a complete classroom lecture note by evaluating the lesson materials and teacher voice input.

The objectives of the research study have been arrived subsequently using the NLP and image processing technologies. Here the content summarizing and specifying the first major task. The lesson input materials and the lecturer voice input has been processed in order to derive the outcome. The topic modeling has conducted in order to identify the relevant topic describes the processed text. Upon generating this text outcome the research was then tried to identify the relevant images used by the teacher to explain some lesson content. The image was processed and mapped with the lesson text. The research results were modeled as a software application on order to test and validate. This research application gives a proper approach for the smart class room environment. Though there are some limitations in this application, as a whole it works properly by achieving the main target areas of the solution.

## REFERENCES

- [1] B. C. a. J. Moonen, **Flexible learning in a digital world**, *New York: Routledge*, 2001.
- [2] S. S. a. K. Ramakrishnan, **An Efficient Speech Recognition System**, *Computer Science & Engineering: An International Journal (CSEIJ)*, vol. 3, no. No. 4, August 2013.  
<https://doi.org/10.5121/cseij.2013.3403>
- [3] M. A. Hearst, "**TextTiling: Segmenting Text into Multi-paragraph Subtopic Passages**", *Association for Computational Linguistics*, 1997.
- [4] P. G. A. A. Alem, **Text Segmentation based on Semantic Word Embeddings**, 2015.
- [5] Z. a. A. G. A. Mehrbod, **A Vector Space Model Approach for Searching and Matching Product E-Catalogues**, *Advances in Intelligent Systems and Computing*, 2014.  
[https://doi.org/10.1007/978-3-642-55122-2\\_71](https://doi.org/10.1007/978-3-642-55122-2_71)
- [6] K. G. C. N. M. Deepali, **A Review Paper on Text Summarization**, *International Journal of Advanced Research in Computer and Communication Engineering*, vol. vol. 5, no. no. 3, 2016.
- [7] L. B. N. Andhale, **An Overview of Text Summarization Techniques**, *International Conference on Computing Communication Control and Automation*, 2016.  
<https://doi.org/10.1109/ICCUBEA.2016.7860024>
- [8] G. S. L. V. Gupta, **A Survey of Text Summarization Extractive Techniques**, *Journal of emerging technologies in web intelligence*, 2, 3, 2010.
- [9] D. S. C. N. Moratanch, **A Survey on Abstractive Text Summarization**, *International Conference on Circuit, Power and Computing Technologies [ICCPCT]*, 2016.  
<https://doi.org/10.1109/ICCPCT.2016.7530193>
- [10] L. Bing, **"Abstractive Multi-Document Summarization via Phrase Selection and Merging"**, 7th International Joint Conference on Natural Language Processing, p. 1587–1597, 2015.  
<https://doi.org/10.3115/v1/P15-1153>
- [11] H. T. N. D. C. Y. L. Wee Meng Soon, **A Machine Learning Approach to Coreference Resolution of Noun Phrases**, in *Association for Computational Linguistics*, 2001.
- [12] F. Y. O. C. J. J. H. Misra, **Text segmentation: A Topic Modeling Perspective**, *Journal of Information Processing and Management*, vol. 47, pp. 528-548, 2011.
- [13] J. W. i. Yin, **A Dirichlet Multinomial Mixture Model-based Approach for Short Text Clustering**, "*Tsinghua National Laboratory for Information Science and Technology (TNList)*", p. 24–27, 2014.
- [14] W. L. a. A. McCallum, **DAG-structured mixture models of topic correlations**, "*international conference on Machine learning, Pittsburgh, Pennsylvania, USA*", 2006.
- [15] M. S. E. K. M. S. H. A. F. A. Kembhavi, **A Diagram is worth a Dozen images**, *ECCV*, 2016.
- [16] J. J. J. a. D. W. Conrath, **Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy**, *International Conference Research on Computational Linguistics (ROCLING X)*, Taiwan, 1997.
- [17] Sarala D., T. Kanikdaley, S Jogi, R.K. Chaurasiya, **Content-Based Image Retrieval Using Hierarchical Color and Texture Similarity Calculation**, *International Journal of Advanced Trends in Computer Science and Engineering*, 7(2), 2018  
<https://doi.org/10.30534/ijatcse/2018/02722018>
- [18] D.W Yan Lai, K.K Kuok, S.y Gato-Trinidad, D.K.X. Ling, **A Study on Sequential K-Nearest Neighbor(SKNN) Imputation for Treating Missing Rainfall**, *International Journal of Advanced Trends in Computer Science and Engineering*, 8(3), 2019  
<https://doi.org/10.30534/ijatcse/2019/05832019>