International Journal of Control and Automation
Vol. 13, No. 2, (2020), pp. 933-943

Summarization of Speech to Text from Reporter in Police Office with Latent Semantic Analysis (LSA) Method

Dessyanto Boedi Prasetyo[1], Wilis Kaswidjanti[2], Hidayatulah Himawan[3], Affan Hilmi Natsir[4]
Universitas Pembangunan Nasional "Veteran" Yogyakarta, Indonesia
[1]dess@upnyk.ac.id, [2]wilisk@upnyk.ac.id, [3]if.iwan@upnyk.ac.id, [4]alfanhilmy@gmail.com

Abstract

Accepting complaint report was a police service to the community. The writing of reported incident in report construction process was time-consuming. This made long queue emerged. One of the solutions developed was by summarizing voice of the complaint uttered by the reporter. The used summarization method was Latent Semantic Analysis (LSA) that was implemented by using Singular Value Decomposition (SVD) algorithm. In addition, the usage of cross method extraction sentence and Speech to Text from Google Speech API was expected to increase the voice to text conversion accuracy. The testing resulted in f-measure and recall values of 92.3% on the first document, while the biggest precision value was 100% with total of 10 data test. The test was performed by compression rate of 30%. The LSA method used was able to summarize by taking attention to words and sentences similarities. Nevertheless, the summarization result depended on the sentence length.

Keywords: *Summarization, Latent Semantic Analysis, Singular Value Decomposition, Cross Method, Speech to Text*

1. Introduction

The mushroom growth in technology motivates an institution or company to increase its service quality. One of them is police institution which is demanded to always protect and serve all community elements fast and precisely. One of the police functions is to serve the community such as accepting complaint report. The accepted report will be recorded to be a data and then will be followed up. Such report usually contains various information, including the incident description. The service started with the police asked about the problem, and the reporter will describe the incident by bringing documents as evidence. After the problem is discovered, the police will make a report in accordance with what has been reported. Each report takes 10-20 minutes depends on the case. The most time-consuming process is writing the incident description while there are a long queue behind.

Summarization is a text generated from one document or more that states important information from relatively shorter original document [1][2]. One of the methods used to summarize documents is *Natural Language Processing* (NLP), in which it uses automatic summarization method [3][4]. The summarization is based on topic which was obtained from semantic similarity between sentences. The method used is *Latent Semantic Analysis* (LSA) implemented by using *Singular Value Decomposition* (SVD). In order to increase LSA accuracy, cross method is used for sentence extraction after SVD. Result from extraction sentence addition with cross method is better than using the previous LSA method [5].
In a research by Chen, Chang, & Chen [6], they use summarization to summarize a news broadcast voice. The research uses speech recognition established by themselves. The

ISSN: 2005-4297 IJCA

accuracy from speech recognition is not accurate enough with 35% error rate. The use of Speech to text from Google Speech API is expected to increase the accuracy of voice to text conversion. Besides, the LSA use with cross method sentence extraction is expected to increase summarization accuracy. This research aims to establish accuracy generated from LSA (*Latent Semantic Analysis*) in summarizing speech to text result from reporter's voice.

This research results in extraction type instead of abstract type of text summarization. The language used is Indonesian Language to input the summarization. The result of Noise Floor test is no more than -1,6 dB

2. Material and Method

NLP is defined as a theoretical field concerning a computational technic used to analyze and represent written text naturally (human language) on one or more linguistic analysis levels with the purpose to obtain human-like language processing which can be implemented in various fields [7]. One of NLP implementations is summarization.

The process was started by recording the complaint voice as data input. After data were obtained, speech segmentation was performed. This aimed to split the voice data into several parts to be converted into sentences. Furthermore, speech to text or speech recognition process was performed to convert voice into text data [3]. In the next step in which text summarization, there were four processes namely preprocessing, TF-IDF weighting, Singular Value Decomposition (SVD), and Cross Method. After going through those four processes, summarization result of complaint voice would be acquired.

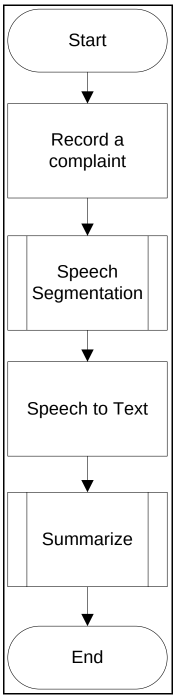
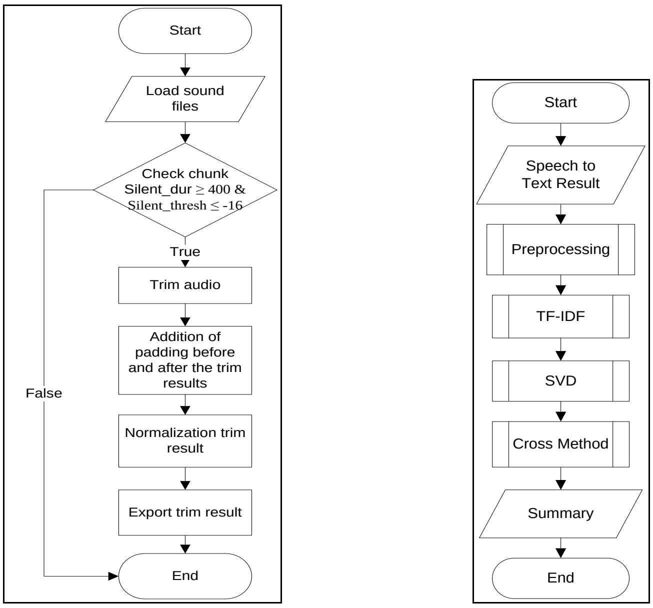


Figure 1. Entire Process Flowchart

Speech segmentation process was a process to split a voice record into several parts (figure 2a). The first process was loading voice record file, then check audio chunks with minimal silence duration of 400ms and maximal silence thresh of -16. If those requirements comply, then the audio was trimmed into audio chunk. Afterwards, silence padding was added before and after chunk as well as normalize such chunk. The last process was exporting the audio chunk to obtain useful audio chunks in voice to text conversion and summarization process.

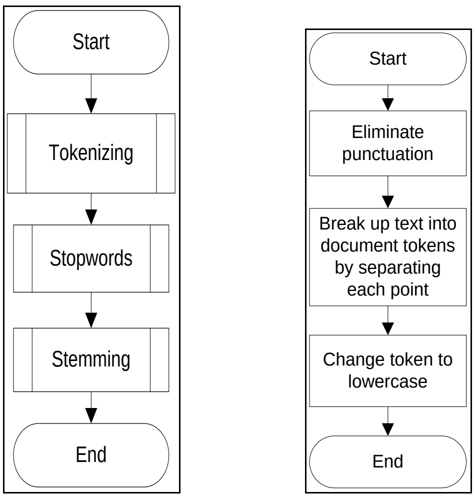
Summarization process was started by taking data from speech to text result. The text which will be summarized was started with preprocessing. Furthermore, weighting was calculated by using TF-IDF. From the weighting result, it was performed calculation of reduction with LSA and SVD in order to clear noise that can disturb the further processes. Based on LSA result, text selection was done by carrying out cross method process. Steps in preprocessing were: parsing, tokenization, filtering, and stemming [8] [3] (figure 3b).



a. Speech Segmentation Flowchart **b. Summary Flowchart**

Figure 2. Flowchart of Speech Segmentation and Summarization Processes

Sub process of preprocessing included tokenization, stopwords deletion, and stemming (figure 3a). This process was a process to make raw text data into ready-to-use data for text analysis process.



a. Preprocessing Flowchart **b. Tokenizing Flowchart**

Figure 3. Flowchart of Preprocessing sub process and Tokenizing

ISSN: 2005-4297 IJCA

Tokenizing sub process depicted process to break up text (picture 3b). The initial part was eliminating the punctuation. After the punctuation was eliminated, the text will be cut if there was period. The next process was change it to lowercase.

Stopwords elimination sub process functioned to delete useless words [9] (figure 4). Result of tokenizing as input, then check the stoplist, if any words found in stopwords dictionary it will be deleted; otherwise, it will be left as a term for further process.

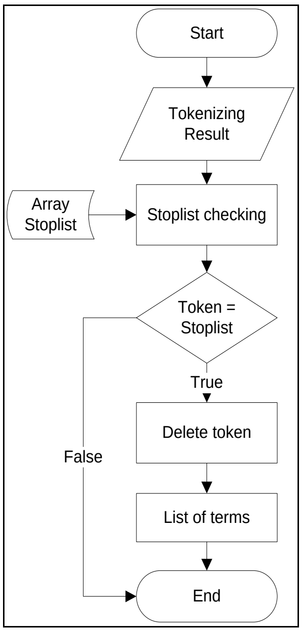


Figure 4. Flowchart of Stopwords sub process

Stemming sub process was a process to find basic word (stem) (figure 5). The process started by inputting the term result after stopwords elimination, then stem checking was done. If the words were in stem list, then it will be left as it was. If it was not found in stem list, infentorial suffix and derrifation suffix as well as prefix and suffix list will be deleted. This process performed by separating all words form either prefix, suffix, or combination of prefix and suffix (confix) into a basic word (*stem*) [10]. When the process was done, the result will generate list of stem.

ISSN: 2005-4297 IJCA

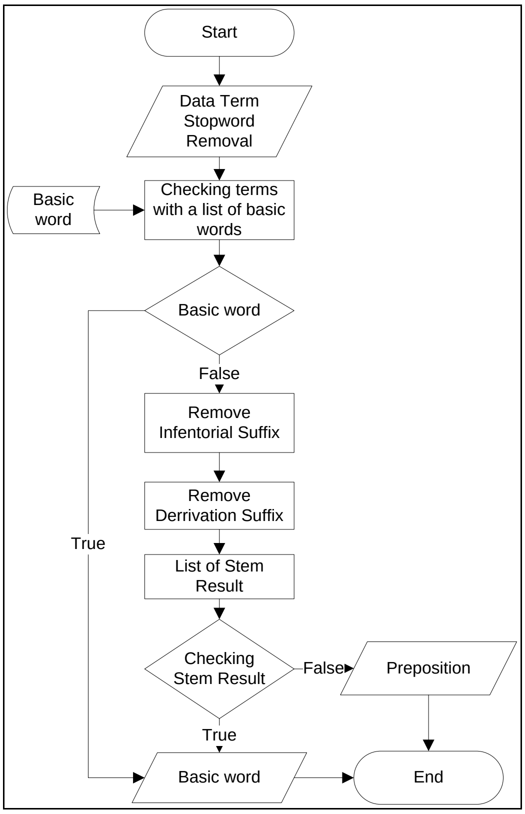


Figure 5. Flowchart of Stemming sub process

Sub process of TF-IDF calculation was weight calculation process of each word (figure 6). The process was started with word generated from stemming, calculated its occurrence frequency on each document, calculated total of document, and enter it to weighting formula, then the result of calculation saved in array. Term Frequency (TF) was occurrence frequency of word in each document. Document Frequency (DF) was total of occurrence word document. D (*Document*) was total of all documents. Inverse Document Frequency (IDF) was calculation result with formula (1) continued by weight calculation with formula (2).

$$IDF = \log {}^{D}/_{df} \qquad (1)$$

$$W = TF * (IDF + 1) \qquad (2)$$

Latent Semantic Analysis (LSA) according to language consists of several important words, in which latent and semantic. Latent means hidden or something which is not yet visible, while semantic comes from Greek word 'semanticos' which means giving important sign, or linguistic branch that studies meaning of a language, code, or other representation code.

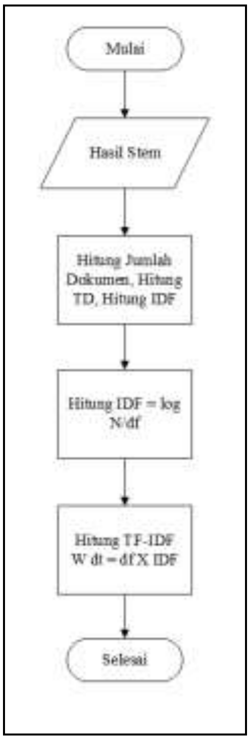


Figure 6. Flowchart of TF-IDF Sub process

Information: Start → Stem Result → Calculate total of documents, TD, and IDF → Calculate IDF = log N/df → Calculate TF-IDF W dt = dfX IDF → finish

LSA elaborates or analyzes the hidden meaning of a language, code, or other representation types in order to obtain important information [8]. LSA is a method based on calculation to extract and represent contextual meaning or word and sentence similarities [11][12]. Word and sentence similarities can be obtained by using *Singular Value Decomposition* (SVD), in which SVD has capacity to reduce noise so that it can increase accuracy on the summarization [2]. SVD calculation sub process was started by input in form of matrix of TF-IDF calculation (figure 7). SVD was a very popular matrix factorization [13].

$$A_{mxn} = U_{mxm} S_{mxn} V_{nxn}^T \qquad (3)$$

Matrix U and V were orthogonal matrixes, while S was diagonal matrix with positive matrix element or zero. Such value was known as *singular value* and matrix U and V were known as related *vector singular* [2].

SVD calculation process can be understood with the following steps:

Given a matrix $A_{mxn} = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$

1. Define Matrix B, if m<=n so $B=AA^T$ if m>n so $B=A^TA$. Due to size of matrix A is 2x2 so m=n then uses formula $B= AA^T$.
   The calculation result is
   $B = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} * \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$
2. Determine the value of eigenvalue B with
   $|\lambda I - B| = \det \begin{bmatrix} \lambda - 32 & 0 \\ 0 & \lambda - 18 \end{bmatrix} = (\lambda - 32)(\lambda - 18) = 0.$
   The calculation result will show $\lambda 1 = 32$ and $\lambda 2 = 18$

ISSN: 2005-4297 IJCA

3. Determine the singular value of matrix A, in which $\sigma 1 = \sqrt{32}$ and $\sigma 2 = \sqrt{18}$
4. Forming Matrix S with condition if m<n so the form of matrix S is $= \begin{bmatrix} \sigma 1 & 0 & 0 & 0 \\ 0 & \sigma 2 & 0 & 0 \\ 0 & 0 & \sigma m \cdots & 0 \end{bmatrix}$.

   if m>n so the matrix S form is $= \begin{bmatrix} \sigma 1 & 0 & 0 & 0 \\ 0 & \sigma 2 & 0 & 0 \\ 0 & 0 & 0 & \sigma m \end{bmatrix}$.

   if m=n so the matrix S form is $= \begin{bmatrix} \sigma 1 & 0 & 0 \\ 0 & \sigma 2 & 0 \\ 0 & 0 & \sigma m \end{bmatrix}$.

   Based on the example, matrix S which will be formed $\begin{bmatrix} \sqrt{32} & 0 \\ 0 & \sqrt{18} \end{bmatrix}$
5. Obtain matrix U and V

   a. Matrix U column is formed from normalization eigenvector of C,

   $C = (\lambda I - B)x = \begin{bmatrix} \lambda - 32 & 0 \\ 0 & \lambda - 18 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

   $\lambda = 32 \rightarrow \begin{bmatrix} 32 - 32 & 0 \\ 0 & 32 - 18 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \end{bmatrix} = 0 \rightarrow U_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

   $\lambda = 18 \rightarrow \begin{bmatrix} 18 - 32 & 0 \\ 0 & 18 - 18 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \end{bmatrix} = 0 \rightarrow U_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rightarrow U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

   b.. Mattrix V column is formed from normalization eigenvector of matrix B

   $D = A^T A = \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix}$.

   Equation of matrix D character

   $|\lambda I - D| = \det \begin{bmatrix} \lambda - 25 & -7 \\ -7 & \lambda - 25 \end{bmatrix}$

   $= (\lambda - 32)(\lambda - 18) = 0.$

   Eigenvalue of matrix D $\lambda 1 = 32$ & $\lambda 2 = 18$.

   Homogen system of matrix D

   $(\lambda I - D)x = \begin{bmatrix} \lambda - 25 & 7 \\ 7 & \lambda - 25 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

   $\lambda = 32 \rightarrow \begin{bmatrix} 32 - 25 & 7 \\ 7 & 32 - 25 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \end{bmatrix} = 0 \rightarrow x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow V_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

   $\lambda = 18 \rightarrow \begin{bmatrix} 18 - 25 & 7 \\ 7 & 18 - 25 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \end{bmatrix} = 0 \rightarrow x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow V_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

   $\rightarrow V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

6. SVD of matrix $A = USV^T$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{32} & 0 \\ 0 & \sqrt{18} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$$

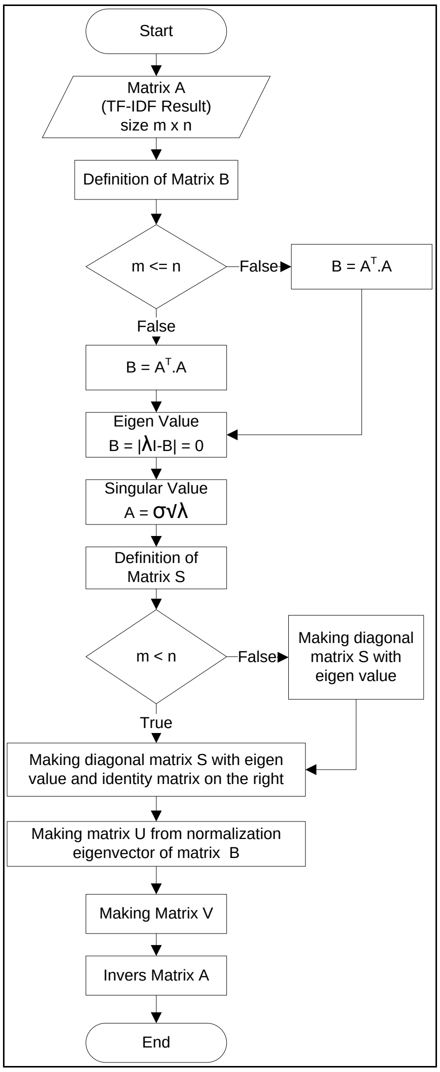


Figure 7. Flowchart of SVD Sub process

Cross method process was further process of SVD calculation (figure 8). Result of matrix SVD especially matrix V^T and matrix S will be processed in this step. The initial part was looking for average value of each line in matrix V^T. Then, check if the average value of matrix is bigger than value of word so the value of word has constant value. Otherwise, if the condition is not fulfilled then the value of word will be made into zero.

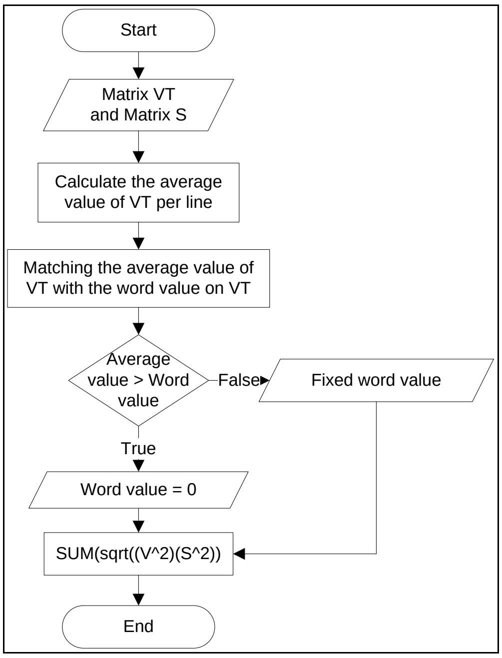


Figure 8. Flowchart of Cross Method Sub process

Cross Method was one of development methods of sentence extraction after SVD calculation process, in other words it was to determine selected sentence from value of matrix V^T [14]. The steps performed were looking for average value of each line of matrix V^T, after the average value was obtained it will be compared with value of word of matrix V^T, if the value of word smaller than matrix value it will be zero, while if otherwise the value of word has constant value. If those two steps have been performed then calculation is carried out with formula:

$$length = \sqrt{V^2 * S^2} \qquad (4)$$

After formula calculation, it will be continued with average value of all length calculation. The sentences showed were sentences with value bigger than average.

3. Result

In this research, evaluation process that will be used is intrinsic evaluation process by using precision method, recall and f -measure. F – measure value was obtained based on value of precision and recall. This evaluation method was the frequently used method in summarization result evaluation process. Method of intrinsic, precision, and recall were used to measure system summarization quality by comparing system and manual summarization (man-made). Precision was the level of summarization accuracy generated from automatic text summarization while recall was level of summarization success [15][16].

$$Precission = \frac{Relevant\ sentence}{\sum System\ Summarization} \qquad (5)$$

$$Recall = \frac{Relevant\ sentence}{\sum Manual\ Summarization} \qquad (6)$$

$$f - measure = \frac{2 * Recall * Precission}{(Precission + Recall)} \qquad (7)$$

System testing will be done on the summarization accuracy. This testing was done on 10 documents. All those documents were calculated its accuracy by comparing manual summarization of police report with compression rate of 60%.

Table 1. Testing Result

Document	Precision	Recall	F-Measure
1	84,8%	78,6%	81,6%
2	100%	75,6%	86,1%
3	99.8%	61,9%	76,4%
4	78,3%	70,7%	74,3%
5	87,6%	63,3%	73,5%
6	99,7%	75,5%	85,9%
7	94,7%	71,6%	81,7%
8	99,7%	71,8%	83,5%
9	96,7%	83,6%	89,6%
10	99,7%	61,1%	75,8%
AVERAGE	**94,1%**	**71,4%**	**80,8%**

The result shows biggest value of f-measure and recall are 89,6% and 83,6% on the ninth document, while the biggest precision value is 96,7%. Average value of f-measure, precision, and recall are 80,8%, 94,1% and 71,4%, respectively.

Summarization testing was carried out with intrinsic test, where precision and recall are used to measure system summarization quality by comparing it with manual summarization. Precision was the level of summarization accuracy, while recall was the level of success of generated summarization. Value of f-measure was used for calculating summarization accuracy.

4. Conclusion

Based on research performed, conclusion can be drawn:

1. Summarization accuracy increase by Latent Semantic Analysis (LSA) from speech to text result can be conducted by using Google API speech to text and sentence extraction by using cross method.

The testing result showed that the biggest value of f-measure and recall was 92.3% on the first document, while the biggest precision value was 100% with 10 test data. This testing was done by compression rate of 60%. The LSA method used can summarize by paying attention to word and sentence similarities. Nonetheless, the result of summarization really depended on the length of each sentence.

References

[1] S. Liu, K. Chen, B. Chen, H. Wang, and H. Yen, "Positional Language Modeling for Extractive Broadcast News Speech Summarization Institute of Information Science , Academia Sinica , Taiwan National Taiwan Normal University , Taiwan," no. Lm, pp. 2729–2733, 2015.

[2] S. K. Bharti, K. S. Babu, and S. K. Jena, "Automatic Keyword Extraction for Text Summarization : A Survey," 2017.

[3] M. Rott and P. Cerva, "Speech-to-Text Summarization Using Automatic Phrase Extraction from Recognized Text," pp. 101–108, 2016.

[4] A. J. Tixier, M. R. Hallowell, B. Rajagopalan, and D. Bowman, "Automation in Construction Automated content analysis for construction safety : A natural language processing system to extract precursors and outcomes from unstructured injury reports," *Autom. Constr.*, vol. 62, no. 2016, pp. 45–56, 2020.

[5] S. S. Lwin, "Extractive Summarization for Myanmar Language," *2018 Int. Jt. Symp. Artif. Intell. Nat. Lang. Process.*, pp. 1–6, 2018.

[6] B. Chen, H. Chang, and K. Chen, "Sentence modeling for extractive speech summarization," *IEEE Int. Conf. Multimed. Expo*, 2013.

[7] J. Hirschberg and C. D. Manning, "Advances in natural language processing," *Science (80-.).*, vol. 349, no. 6245, 2015.

[8] F.-A. J. Oi-Mean Foong, Suet-Peng Yong, "Text Summarization Using Latent Semantic Analysis Model in Mobile android Platform," *9th Asia Model. Symp.*, 2015.

[9] L. Cagliero, P. Garza, E. Baralis, and P. Torino, "ELSA : A Multilingual Document Summarization Algorithm Based on Frequent Itemsets and Latent Semantic Analysis," vol. 37, no. 2, pp. 1–33, 2019.

[10] A. S. Nayak, A. P. Kanive, N. Chandavekar, and R. Balasubramani, "Survey on Pre-Processing Techniques for Text Mining Survey on Pre-Processing Techniques for Text Mining," no. September, 2016.

[11] S. R. Vrana, D. T. Vrana, L. A. Penner, S. Eggly, and R. B. Slatcher, "Social Science & Medicine Latent Semantic Analysis : A new measure of patient-physician communication," *Soc. Sci. Med.*, vol. 198, no. December 2017, pp. 22–26, 2018.

[12] P. J. Kwantes, N. Derbentseva, Q. Lam, O. Vartanian, and H. H. C. Marmurek, "Assessing the Big Five personality traits with latent semantic analysis," *Pers. Individ. Dif.*, vol. 102, pp. 229–233, 2016.

[13] C. D. Martin and M. A. Porter, "The Extraordinary SVD," *Am. Math. Mon.*, vol. 119, no. 10, pp. 838–851, Dec. 2012.

[14] D. K, Geetha J, "Kannada Text Summarization," *Int. Conf. Adv. Comput. Commun. Informatics (ICACCI*, pp. 1508–1512, 2015.

[15] R. Leaman, R. Khare, and Z. Lu, "Challenges in clinical natural language processing for automated disorder normalization," *J. Biomed. Inform.*, vol. 57, pp. 28–37, 2015.

[16] M. Sevenster, J. Buurman, P. Liu, J. F. Peters, and P. J. Chang, "Natural Language Processing Techniques for Extracting and Categorizing Finding Measurements in Narrative Radiology Reports," pp. 600–610, 2015.

ISSN: 2005-4297 IJCA