

An Approach to Develop a Framework to Enhance the Performance of Digital Notes Based on Auto Arranger

Noopur Srivastava¹, Bineet Kumar Gupta², Neeraj Kumar Tiwari³

¹M.Tech (CS) Department of Computer Science and Engineering from Shri Ram Swaroop Memorial University, Deva Road, Barabanki-225003, Uttar Pradesh India.

²Assistant Professor Faculty of Computer Science and Engineering, Shri Ram Swaroop Memorial University, Deva Road, Barabanki-225003, Uttar Pradesh India.

³Assistant Professor Faculty of Computer Science and Engineering, Shri Ram Swaroop Memorial University, Deva Road, Barabanki-225003, Uttar Pradesh India

Abstract- In the past decade provision for preparing notes were in single copy for all subjects and finally writes notes in different subjects copy. Now in the advance age of technology and viewing the future aspects of students, it is well demanded to prepare their notes digitally and distribute them automatically in their respective subjects and no need to prepare separate notes manually. For better distribution process we need a reliable and efficient distribution approach by which one note automatically distributed in their respective subjects. In our approach part 'cue word' are those relative words of a prescribed subject that helps to distribute a single note into its respective folders on the basis of frequency of cue words. Proposed approach will announced in different sectors like letter writing, student's notes making, script managing, news reporting and many more criteria related to writing work and managing prepared notes.

Keywords: Data mining, data warehouse, artificial intelligence, SSA, Digital Notes.

I. INTRODUCTION

An approach to introduce the essential research area of the data mining, algorithm implementation presented in the paper, where, analyses of one subject note and bifurcate it into its respective subject on the basis of 'cue words' were made. Looking the future aspects, the algorithms of Source Separation Algorithm (SSA) are implemented. Key to algorithms are exploiting detailed source models, using different filtering ideas to distribute the single file, and employing explicit models for background and file division. An attempt to demonstrate strong performance which is robust to distribution of single note file presented [8]. Good scaling properties are obtained using Artificial Intelligence Technique.

An artificial intelligent approach fetches the page from the notes and store in suitable folder on the basis of special type of survey after collecting the subject name. This method is based on artificial intelligence (AI) techniques and implemented to improve the interpretation accuracy in separation of notes according to subject using implementation of SSA (Source Separation Algorithm) with faster manner [8], [11].

AI techniques are applied to establish classification features for Subject verification based on the collected data known as cue word. The features are applied as input data to fuzzy logic, artificial intelligence, data warehouse, data mining to diagnose accurate subject of notes. The experimental data of stop word is used to evaluate the performance of proposed method as in the form of separation. The results of the various evaluating methods are classified using AI techniques and compared with the extraction of stop words and frequency of cue word within the paragraph.

II. THE ASPECTS OF AUTO ARRANGER

Analysis of actual note is based on data research of subject. All research is nothing but it is the huge collections of combine experience of cue words that is already taken earlier in the form of frequency of best suitable words which are really actual words for that relevant subject [2], [12], [10].

The experiences gathered are unified all cue words and put them in a model which separate the digital notes into different subject by using the approach of source separation algorithm. Simple demonstration of auto arranger shows the extraction of digital notes for the process of separation in the form of different subject, which are stored in the database (Fig. 1). Location of database could be considered in centralized or distributed environment

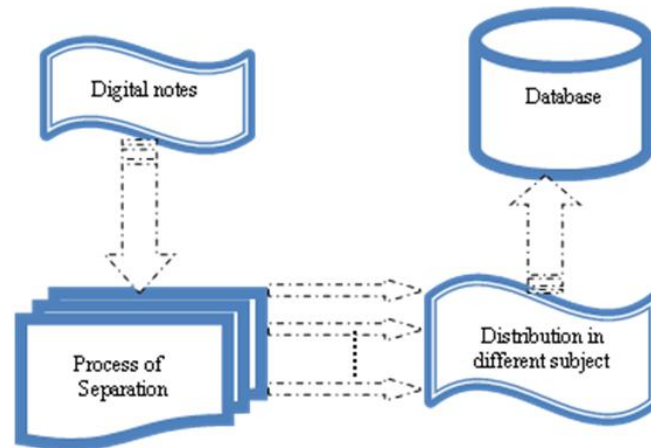


Fig. 1: Process of Auto Arranger

A. Artificial Intelligence Approach

This technology is based on fuzzy logic in which artificial intelligence tool also works. Here artificial intelligence helps to maintain the discipline calculation of percentage which signifies the actual subject percentage. While, fuzzy logic helps in percentage frequency calculation [10], [13]. In combination of both approaches a unique formula for the word prepared, which, helps the user to identify more correctly without wasting time and money.

B. Data Warehouse Approach

It is a central repository of data which is created by integrating data from one or more disparate sources. Data warehouses store current as well as historical data and are used for creating trending reports for senior management reporting such as annual and quarterly comparisons. By this approach, proposed tool collect the cue words and store high frequency word in respective subject at the time of adding paragraph of new subject. This process is going on for future reference when these words will match by auto arranger for final distribution of single note into different subject’s collection.

Collection or summarization of cue words after removing the stop words, worked as a Data ware house are presented in present piece of work (Fig. 2).

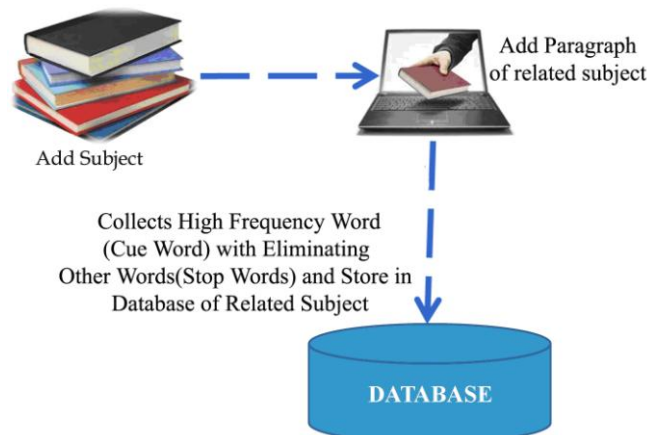


Fig. 2: Data warehouse of Digital Notes

C. Data Mining Approach

The goal of the data mining process is to extract information from a data set and transform into an understandable structure for further use.

In present attempt, comparison of cue words with new attached notes and make frequency table is worked as Data Mining. Here the role of Data mining is to make frequency of matched cue words and find how many times the subject word is coming in the compared text. By this approach this tool works more efficiently [3], [7].

III. PROPOSED TECHNIQUE

An attempt to generate Artificial Intelligence where everything can be automatically operated made in present piece of work [5]. In this technique attempt were made to present one step more that is automated text arrangement which automatically arrange the text in different subjects without wasting the time in arranging in their respective subject manually. The adopted aspect automatically breaks down the single notes into different parts according to criteria. In proposed help line user may concentrate only writing their work in place of managing their work.

The proposed model helps student, notes maker, news manager, speech designer and many more persons for managing their written work in more refine manner and save their time in converting their rough notes into fair notes[5], [9]. Considering the facts that “Save Paper, Save Tree, Save Environment, Save World”, present result will help the people to prepare their notes digitally and no need to maintain the notes subject wise. This model is very efficient to understand or recognize that for which subject it is written. These are all basic reasons to announce this model as helpful means for managing the notes in future with saving precious time.

Impact of Proposed technique:

If auto arranger is not used, then also data still store in data structure but it is difficult to separate the data in single area as it is stored in bulk and further separating the data is time talking as well as difficult too. Whereas if auto arranger is used it will automatically separate the data as per user choice in their specific area. Hence we can say that our proposed work is a beneficial for user to have

- Ease of use
- Effective information retrieval
- Reduce time complexity
- Handy for arrangement of more storage of data

A. SSA Algorithm Approach

Source Separation is a tempting appliance for several audio applications. This appliance is used as a pre-processing technique as well as post -processing technique. Pre-processing technique is used for hearing aids on the other hand post-processing technique is used for de mixing of stereo recording. An early approach which uses implementation of source separation algorithm is EM Model [8], [14]. EM model is a sound model which is used to separate the signal into two different estimates. This is a concept by which SSA algorithm bifurcate the original voice into male voice and female voice [14].

Now in this area, a new model is implemented by using SSA Algorithm i.e., Auto arranger a text model. It is based on extracting a single note text into different subject folder automatically. It distributes single file into respective folder on the basis of frequency of cue words.

This investigation described a multitude of methods for source separation, employing a wide range of mining learning techniques as well as knowledge of separation of digital notes by which one can save the notes. This algorithm first selects the text in the form of digital notes which works as input for the auto arranger. Auto arranger divides or separates the notes with respect to the desired subject with the help of the cue words [1], [10], [13]. An important lesson to note that, although the source separation problem can be formulated in very general terms, the solution cannot. The formula to implementation of SSA algorithm is given below.

$$SN_j = (ON - Sw_k) + \sum (Fw_1 + w_2 + \dots + w_i) / n$$

Where

SN_j- After separation of digital note, it is divided into different notes according to their respective subject as output form and j is number of division.

ON- It is that digital note that takes as input form. Here the cardinality is “one too many” where one digital note can be broken down into different notes according to their respective subjects without wasting time and reliable to convert notes automatically arrange in their proper place that like Artificial Intelligence Technology.

Sw_k- It is number of “Stop Words” that are hard coded in database. K represents number of stop word store in our database. Example- is, the, go ... etc.

$\sum (Fw_i)$ - Here ‘w’ represents Cue Words that is important word of relative subject. Here ‘i’ is number of cue words. ‘F’ shows frequency means number of time a particular word has come in the original text. \sum is addition of all frequency of all cue words situated in text.

n- Here n is total word comes in original text.



Fig. 3: A framework to enhance data Structure

B. Auto Arranger: Deployment Strategy

Basically all the following researches and data are based on the experience of all the subjects cue words around the world and the following research are nothing but the huge collections of combine experience of cue word collection as shown in Fig. 3.

Steps of Research Algorithm Implementation:

- The first step in the analysis is to abstract over the input, in order to find “stop words” comparison where “stop words” are a part of data warehouse. The different example are ‘a’, ‘an’, ‘the’.
- The second step is to abstract over the implementation, to find “cue words” comparison that is actual words stored for future comparison by external source file.
- For the summation program, this could be just the number of lines of code executed, or it could be more detailed, measuring the number of comparison, executed by the algorithm.

IV. CONCLUSION

Proposed model is more efficient and more reliable to arrange a single file into different folders according to their subjects than manual system. This model is very helpful for those students who prepare notes in a single document and want to arrange in different folder according to their subject. It also enhances the performance of data Structure. Whole model is depending on the cue word.

Presented cue word is a key of whole implementation that can reduce the maintaining effort and increase efficiency of data marinating capabilities. Due to the use of cue word present tool is more efficient in aspects of data structure. It is also speculated that these current capabilities can grow to support future notes making strategies, where human intelligence stops and utilized their valuable time more efficiently in this age of fast growing era.

ACKNOWLEDGEMENT

This study was a part of my M.Tech. Final year dissertation conducted at the Faculty of Computer Science and Engineering, Shri Ramswaroop Memorial University, Lucknow Deva Road, Uttar Pradesh, India

REFERENCES

- [1] J. Anemuller, “Convolutional Blind Source Separation”, Ph.D. thesis, University of Oldenburg, Germany, 2001.
- [2] Araki S., “Limitation of frequency domain blind source separation for convolutional mixture of speech”, *proc IEEE HSC-01*, pp. 91-94, 2001.
- [3] H. Attias, “Independent Factor Analysis Neural Computation”, 11, pp. 803-851, 2009.
- [4] M. Baeak, and U. Zolzer, “Performance Analysis of a Source Separation Algorithm”, *DAFx-02*, Hamburg, Germany, pp. 207-210, 2002.
- [5] J.A. Bell, and J.T. Sejnowski, “An information maximization approach to blind separation and blind deconvolution. *Neural Computation* 7”, pp. 1129-1159, 1995.
- [6] Y.Ephraim, “Statistical model based speech enhancement system”, *proc IEEE* 80(10), 1526-1555, 1992.
- [7] B.Gupta, and md. Hussain, “Algorithm to Evaluate the Rank of Research Papers Using Citation Graph” in *CiiT International Journal of Data Mining Knowledge Engineering*, ISSN NO 0974, April Issue 2012 Paper ID 3033
- [8] A. Hagai, “New EM algorithms for source separation and deconvolution with a microphone array”, *IEEE Transaction*, 2008.
- [9] R.K. Olsson. And L.K. Hansen, “Blind Separation of More Sources than Sensors in Convolutional Mixtures”, *International Conference on Acoustics on Speech and Signal Processing*, vol. 5, pp. 657-660, 2006.
- [10] B. Sharef, N. Omar, and Sharef Zeyad., “Automated Arabic Text Categorization based on the frequency ratio accumulation”, *The international arab Journal of Information Technology*, 11(2), 2014.
- [11] H.N. Thi, and C. Jutter, “Blind source separation for convolutional mixtures”, *Signal processing*, 45:209-229, 1995.
- [12] T. Virtanen, “Sound Source Separation in Monaural Music Signals”, 2006.
- [13] O. Yilmaz, and S. Rickard, “Blind Separation of Speech Mixtures via Time-Frequency Masking”, *IEEE Transactions on Signal*, 2002.
- [14] Matthias Baeck, and Udo Zolzer, “Real-Time Implementation of a Source Separation Algorithm”, 6th international conference on digital audio effects (DAFx-03), London, UK, 2003.