



## A Novel Framework for Multi-Document Temporal Summarization (MDTS)

Kishore Kumar Mamidala <sup>1\*</sup>, Suresh Kumar Sanampudi <sup>2</sup>

<sup>1</sup> Associate Professor, Department of Computer Science & Engineering, Vivekananda Institute of Technology & Science, Karimnagar, India.

<sup>2</sup> Assistant Professor & HOD, Department of Information Technology, JNTUH College of Engineering Jagtial, Telangana, India.

### Abstract

Internet or Web consists of a massive amount of information, handling which is a tedious task. Summarization plays a crucial role in extracting or abstracting key content from multiple sources with its meaning contained, thereby reducing the complexity in handling the information. Multi-document summarization gives the gist of the content collected from multiple documents. Temporal summarization concentrates on temporally related events. This paper proposes a Multi-Document Temporal Summarization (MDTS) technique that generates the summary based on temporally related events extracted from multiple documents. This technique extracts the events with the time stamp. TIMEML standards tags are used in extracting events and times. These event-times are stored in a structured database form for easier operations. Sentence ranking methods are build based on the frequency of events occurrences in the sentence. Sentence similarity measures are computed to eliminate the redundant sentences in an extracted summary. Depending on the required summary length, top-ranked sentences are selected to form the summary. Experiments are conducted on DUC 2006 and DUC 2007 data set that was released for multi-document summarization task. The extracted summaries are evaluated using ROUGE to determine precision, recall and F measure of generated summaries. The performance of the proposed method is compared with particle swarm optimization-based algorithm (PSOS), Cat swarm optimization-based summarization (CSOS), Cuckoo Search based multi-document summarization (MDSCSA). It is found that the performance of MDTS is better when compared with other methods.

### Keywords:

Natural Language;  
Processing;  
Text Summarization;  
Multi-document Summarization;  
Extractive Summary;  
Temporal Summarization.

### Article History:

<b>Received:</b>	25	January	2021
<b>Revised:</b>	04	March	2021
<b>Accepted:</b>	09	March	2021
<b>Published:</b>	01	April	2021

## 1- Introduction

The information is overloaded with redundancy, and enormous types of content spread unevenly over the documents. So, for a user, difficulties arise in finding appropriate content from such large amounts of data. The user satisfaction level is low due to the lack of efficient search engines. Summarization is the process in which the data is presented in a brief and precise way.

Summarization is abstracting or extracting key content from one or more information sources. This process generally identifies the vital information in a document or set of related documents and presents in condensed form. Different users may have different backgrounds and preferences in writing their summaries. So, the meaning of vital information and redundancy differs from user to user [2]. Temporal summarization includes efficient monitoring and handling of information associated with an event over time. For example, in online news articles, the same topic is published with different views. Summarizing the information from these multiple sources may contain the redundant data into summary. The happenings in the news are associated with time information. So, the need raises to build a Multi-Document Temporal Summarization (MDTS) which generates the summary with vital and non-redundant information.

\* **CONTACT:** Kishore.mamidala@gmail.com

**DOI:** <http://dx.doi.org/10.28991/esj-2021-01268>

© 2021 by the authors. Licensee ESJ, Italy. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Events are things that happen or occur in the world (like weddings, birthdays, and parties...) at a certain time or over a given time and in a given place. They are typically dynamic occurrences with causes and effects, a clear beginning and end, and bring about some perceptible change in the world [1].

A Temporal Expression can be defined as any element in language that lexicalizes the concept of time in terms of recognized and generally quantifiable temporal units. Temporal expressions are classified into four categories according to the types of instants or intervals they describe: dates, times, durations, and sets [3, 10, 17]. The events and times which are temporal entities are the basis of any MDTs. The text is annotated by the identification and classification of different events and times from the documents provided [5]. TIMEML [6, 7] is an annotation tool in which event and time tags are presented. This tags syntax is used to extract and represent the events and times from the documents.

This paper presents a novel framework for temporal summarization on multiple documents. Event based summarization method is developed in the proposed framework. Sentence scoring and sentence selection mechanisms are introduced to generate the efficient summaries. Sentence scoring measures are developed based on the presence of events and time in the text. Similarity Score measure is used to eliminate the redundant sentences from that of extracted. Then top-ranked N number of sentences are selected as a summary based on the required summary length. The section 2 of the paper explains the related work in text summarization area. The framework of MDTs is explained in section 3. Experimentation, and the evaluation results are detailed in section 4. And lastly, the conclusion and future scope is discussed in section 5.

## 2- Literature Review

Depending on the input documents the summarization methods are classified into single or multi-document summarization. If a single document is given as input for summary extracted, such a method is referred to as single document text summarization [2, 15]. If more than one document is given as input to generate the summary, such a method is referred to as multi-document summarization [13, 18, 20-22]. Temporal summarization generates the summary depending on the temporal events extracted from the text [1]. The scope of research in Temporal Information Processing is broadened, ranging from classical linguistic theories of tense and aspect to current computational approaches for extraction of Temporal Information [11].

James Allen [8] developed a theory, named "Allen's Interval. "Algebra" provides a conceptual model of time that captures the different ways in which eventualities may be related to each other. Allen considers that every temporal expression or event can be presented as a temporal interval having a start point and an endpoint on a timeline. In this concern, Allen defines a set of thirteen fundamental (binary) interval relations, where six are inverses of the other six, excluding equality. Filatova and Hovy [3] also proposed a model for temporal relations representation known as the "Time Stamping of Events." This model serves to arrange the contents of news stories into a timeline. This procedure consists of assigning a calendar time point or interval to all events in a text. However, this model does not capture information in many cases and sometimes loses information or misinterpret the information.

The state-of-art of Temporal Reasoning proposes several constraint-based models [16, 17]. The events in the text forms the nodes of the network and the links forms the temporal relations in which the order of these events occur. Thus, formed network is viewed as temporal constraint satisfaction problem (TCSP). Reasoning is applied for TCSP to derive of unknown temporal relations among the events. This connections of events with time gives better readability for the generated summary. Temporal modelling [4, 9] for COVID-19 data set is performed using the concepts or word2vec and word embeddings for each occurrence of event in a time span. It measures the semantic changes of event over a time. Temporal Information Processing includes two major tasks: one is to automatically recognize and extract temporal entities in the narrative (namely temporal expressions and events), and the other is to discover temporal relations between these entities and inferring the type of each recognized one.

Handling temporal annotation requires the development of annotations schemes and the construction of annotated corpora. In this concern, a lot of ongoing research focuses on developing annotation schemes to extract, model, and interpret temporal information in natural language texts. TimeML [14, 5] (Time Markup Language) is the latest annotation scheme. TimeML has been standardized to an ISO international standard for temporal information markup, ISO-TimeML (ISO-2007) [6]. XML-based annotation format are used to define <Event>, <TIMEX> and <TLINK> tags. <Event> tag annotate the events, <TIMEX3> annotate the time expressions, and <TLINK> tag identifies the temporal link among the tagged events.

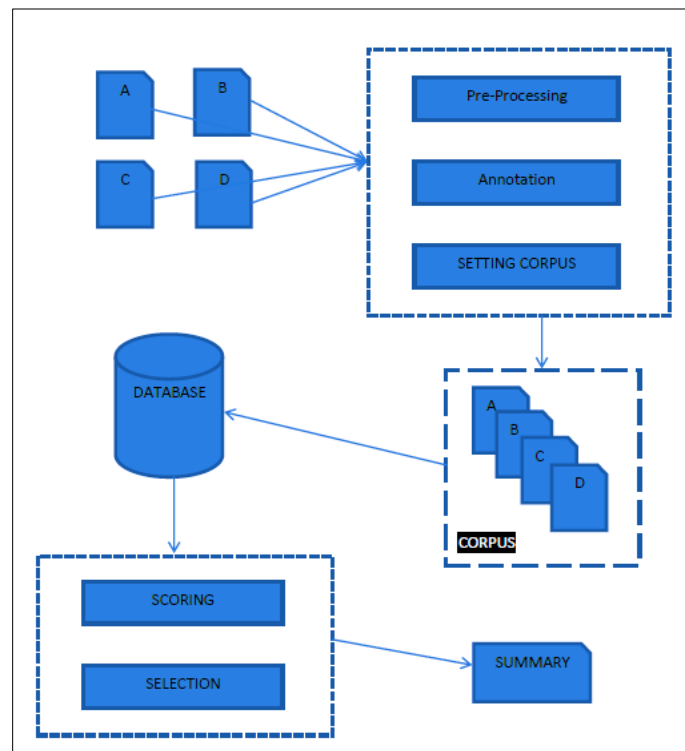
## 3- A Framework for Multi Document Temporal Summarization (MDTS)

A framework for multi-document temporal summarization (MDTS) is shown in Figure 1. The work is divided into three tasks 1) Temporal information processing in which the events, times, and temporal expressions are extracted from the corpus. 2) The corpus along with extracted data are stored in a database format. 3) Sentence scoring and selections are performed to generate the summaries.

The input to the system is a set of text files. The files may be involving News Reports, Medical Data, Financial sets. The files must contain timed text data i.e. the data which involves time expressions or time values along with the events. TREC has released the datasets for temporal summarization in the tracks organized during 2013, 2014 and 2015.

The pre-processing step involves the elimination of stop words from the text. Then the remaining text is tagged with parts of speech tagger. Event recognition involves the identification of some attributes related to each event like tense, aspect, modality, polarity, and class. <EVENT> tag is used for events extraction and <MAKEISNTANCE> tag for event instances with attributes. <TIMEX3> tags different time expressions such as times, dates, durations. <TLINK> tags are used to establish the event-event and event-time relations among the events and times obtained [19].

All the input files are processed into TimeML format and are set in a corpus. In this step, the corpus is imported to a database. The corpus is initially processed by an XML, which retrieves document level data as well as all temporally annotated information, and places it into a SQLite database. Each and every TimeML tag is made into a table and its attributes as columns. A total of 10 tables are created initially. The tables, namely, documents, sentences, events, timex3, instances, signals, tlinks, alinks, slinks, info, etc.



**Figure 1. A Framework for Multi-Document Temporal Summarization.**

Sentence Scoring is the process where scores are allotted to every sentence of each document. The scores were calculated on different aspects of sentences covering the events and times data available in the documents. Based on these scores, the sentences are ranked. Scoring activities include two types scoring parameters Events based sentence ranking and Time-based sentence ranking.

The event-based ranking of a sentence is termed TSR1, in which the score of a sentence depends on the number of events present in the sentence shown in Equation 1. The greater number of events in a sentence constitutes the importance of that sentence to be present in the final summary. TSR1 is normalized based on corpus mode.

The TSR1 score of a Sentence ‘*i*’ is defined as follows:

$$TSR1(S_i) = \frac{event\_count(s_i)}{\sum_{j=0}^n event\_count(s_j)} \quad (1)$$

Where; ‘ $event\_count(s_i)$ ’ denotes the number of events contained in the sentence ‘*i*’, In Normal Mode, ‘*n*’ denotes the total number of sentences present in all the documents. In Doc-specific Mode, ‘*n*’ is the number of sentences present in single document.

The temporal-based ranking of a sentence is termed TSR2, in which the score of a sentence depends on the number of time expressions present in the sentence shown in Equation 2. The more the number of such time expressions indicates that the sentence is valuable than others. TSR2 is then normalized based on the mode of the corpus.

The TSR2 score of a Sentence ‘*i*’ is defined as follows:

$$TSR2(S_i) = \frac{timex_{count}(s_i)}{\sum_{j=0}^n timex_{count}(s_j)} \quad (2)$$

Where; ' $timex_{count}(s_i)$ ' denotes the number of time expressions associated with the sentence ' $i$ ', In Normal Mode, ' $n$ ' denotes the total number of sentences present in all the documents.

In Doc-specific Mode, ' $n$ ' is the number of sentences present in single document.

Sentence scoring concentrates on the similarity checking among the sentences extracted. Topic Relevance Score (TRS) of a sentence is defined as the score generated by the comparison of similarity or intersection of that one sentence to other sentences in the extracted text shown in equation 3. The less similarity score indicates the importance of that sentence.

The TRS of sentence ' $i$ ' is defined as follows:

$$TSR(S_i) = \frac{len[Sim(s_i, s_o)]}{\left(\frac{len[s_i] + len[s_o]}{2}\right)} \quad (3)$$

Where;  $Sim(s_i, s_o)$  denotes the similarity measure of sentence  $s_i$  with sentence,  $s_o$ ;  $len[ ]$  denotes the length as the number of words, In Normal Mode, ' $n$ ' denotes the total number of sentences present in all the documents. In Doc-specific Mode, ' $n$ ' is the number of sentences present in single document.

The summary generation process is mainly based on Sentence Selection by using extraction methods. Each sentence has assigned a score, and ranking is done based on the sentence scores computed in the previous step. Depending on the length of the summary required appropriate number of sentences are extracted. The top  $n$  sentences are selected to form the summary. The summaries are stored into text files, which makes it easy for the summary evaluation process.

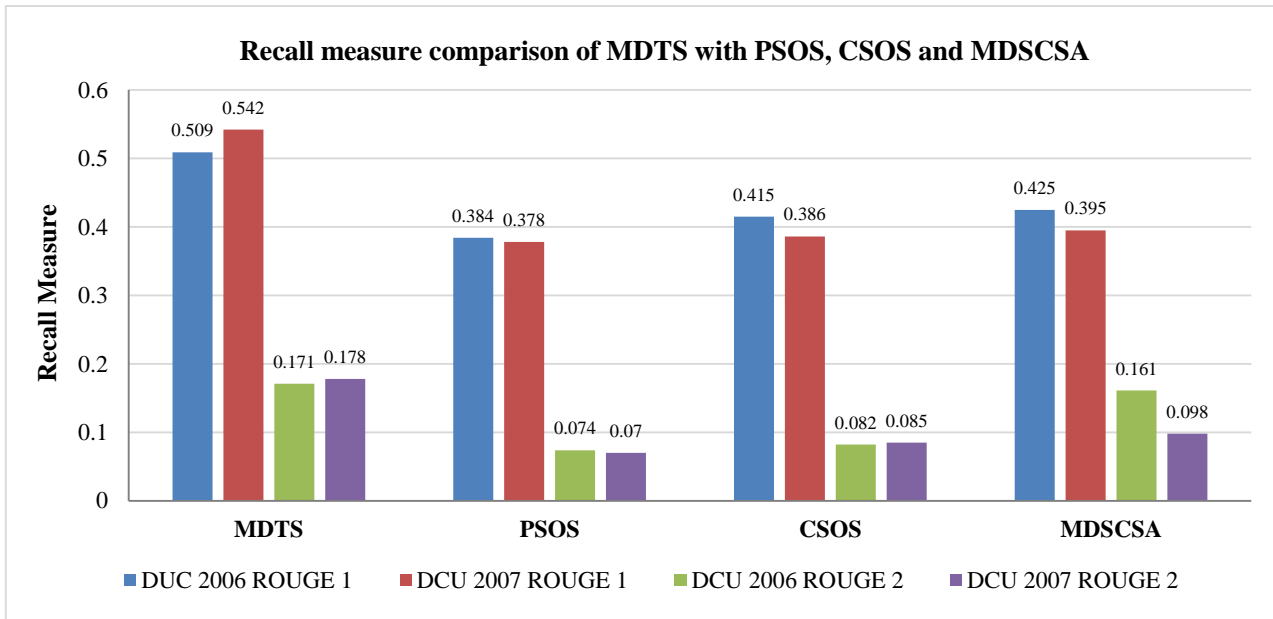
### 3- Results and Discussion

Experiments are conducted on the Document Conference Understanding (DUC) DUC 2006 and DUC 2007 dataset. The data set contain nearly 2400 English articles released for the experimentation in the track [11]. The average number of sentences per document range from 30 to 37, summaries generated are evaluated in for a length of 250 words summaries. Recall-Oriented Understudy Gisting Evaluation (ROUGE) is a standard tool used to evaluate the generated summaries. ROUGE is used to evaluate a summarization system. It requires two types of summaries they are system generated summaries referred to as 'peer summaries' and reference summaries or gold standard summaries known as 'model summaries.' ROUGE can handle any number of peer summaries (if generated by multiple systems) and any number of model summaries. ROUGE N Scores are used to measure the performance of generated summaries. ROUGE 1 and ROUGE 2 scores are used

All the metrics calculate three scores, namely, Recall, Precision, and F-Score. All of the scores are a measure of relevance of sentences extracted in the generated summary compared with that of human summaries. System generated summaries are compared against some reference or model summaries. The most commonly used metrics are ROUGE-N, ROUGE-1, ROUGE-2, all the metrics calculate three scores, namely, Recall, Precision, and F-Score. These scores are calculated for generated summaries and are compared with the existing methods namely PSOS, CSOS, MDSCSA. These scores are calculated for DUC 2006 and DUC 2007 [12] dataset and tabulated in Table 1.

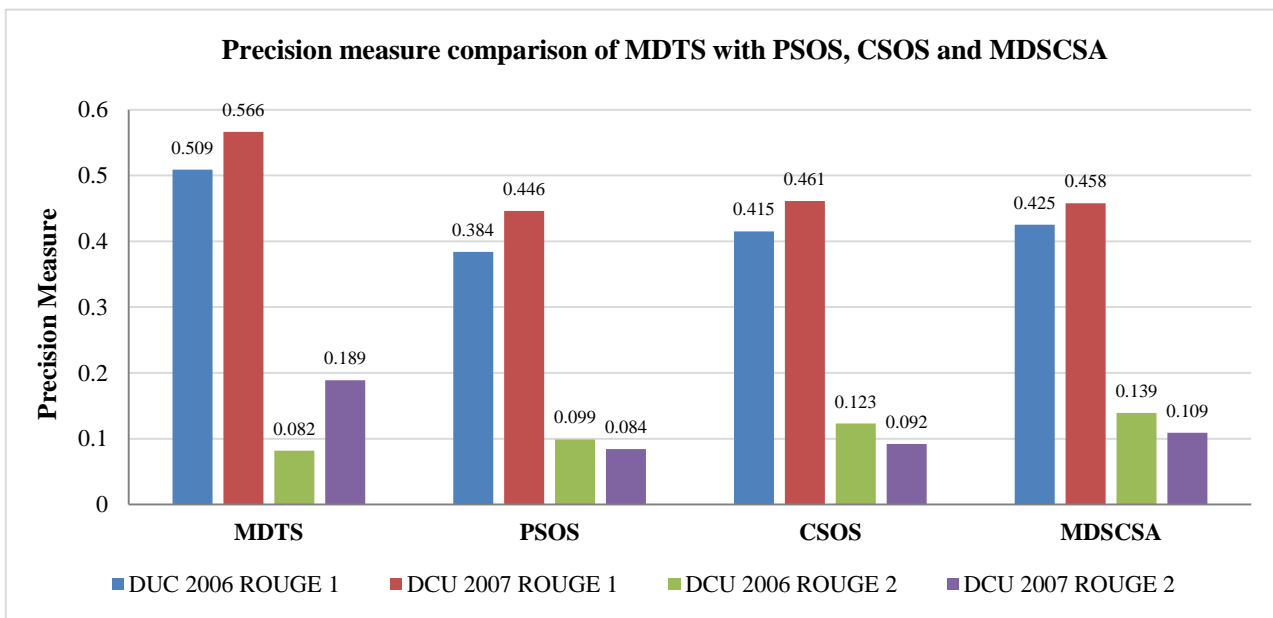
**Table 1. ROUGE N Score calculation for MDTS comparing with PSOS,CSOS and MDSCSA tested on DUC 2006 and DUC 2007 dataset.**

Method	Evaluation Metric	DUC 2006		DUC 2007	
		ROUGE 1	ROUGE 2	ROUGE 1	ROUGE 2
MDTS	Recall	0.509	0.171	0.542	0.178
	Precision	0.522	0.184	0.566	0.189
	F-Score	0.516	0.177	0.558	0.181
PSOS	Recall	0.384	0.074	0.378	0.070
	Precision	0.441	0.082	0.446	0.084
	F-Score	0.411	0.078	0.409	0.076
CSOS	Recall	0.415	0.082	0.386	0.085
	Precision	0.431	0.099	0.461	0.092
	F-Score	0.423	0.090	0.420	0.089
CSMDS	Recall	0.425	0.161	0.395	0.098
	Precision	0.436	0.123	0.458	0.109
	F-Score	0.431	0.139	0.424	0.103

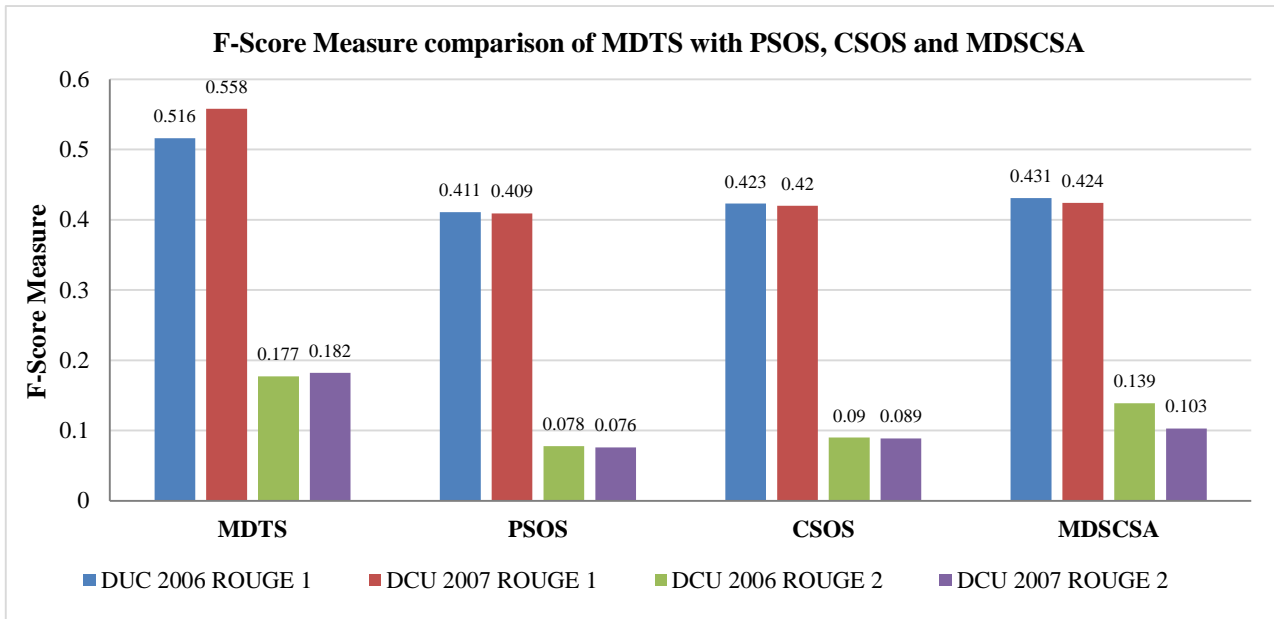


**Figure 2.** Comparison of Recall measure of MDTS with PSOS, CSOS and MDSCSA for DUC 2006, DUC 2007 dataset.

ROUGE 1 and ROUGE 2 scores are calculated for DUC 2006 and DUC 2007 data set and recall comparison is shown in Figure 2. The recall of the proposed method MDTS is found to be better when compared with other method. The Recall value of MDTS is 0.509 for DUC 2006 dataset and 0.542 for DUC 2007 dataset. Similarly, the ROUGE 1 precision score of MDTS is found to be 0.522 for DUC 2006 and 0.566 for DUC 2007 dataset. The precision of the proposed method is compared with other methods and is shown in Figure 3. ROUGE 1, F- score of MDTS is 0.516 for DUC 2006 and 0.558 for DUC 2007 dataset. The comparison of MDTS F-score with other methods is shown in Figure 4. It is observed that high precision and F score of the resulted for MDTS when compared with PSOS, CSOS and MDSCSA. The ROUGE N scores tells how much the system generated summaries are relevant to human generated systems.



**Figure 3.** Comparison of Precision measure of MDTS with PSOS, CSOS and MDSCSA for DUC 2006, DUC 2007 dataset.



**Figure 4.** Comparison of F Score measure of MDTS with PSOS, CSOS and MDSCSA for DUC 2006, DUC 2007 dataset.

## 4- Conclusion

A novel framework for multi-document summarization is presented in this paper. The given documents are pre-processed for annotation. The event, times, and event-time relations are extracted and stored in a database. Depending on the number of events and times present in the sentences. Sentences scores are ranked based on their scores. Redundant sentences are eliminated, and top ranked sentences are selected based to generate summary. DUC 2006 and DUC 2007 datasets are used for experimentation. The generated summaries of various modes are evaluated using the ROUGE tool, in which the recall and precision scores are observed and tabulated. The results of the proposed model have resulted high recall, precision and F score when compared with the existing methods like PSOS, CSOS, MDSCSA methods. The proposed method has retrieved relevant and non-redundant sentences into the summaries which resulted in the highest scores when compared with the existing methods. This framework can be applied to legal documents and medical domain data in future.

## 5- Declarations

### 5-1- Author Contributions

Conceptualization, K.K.M. and S.K.S.; writing—original draft preparation, K.K.M. and S.K.S.; writing—review and editing, K.K.M. and S.K.S. All authors have read and agreed to the published version of the manuscript.

### 5-2- Data Availability Statement

The data presented in this study are available in article.

### 5-3- Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### 5-4- Conflicts of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

## 6- References

- [1] Zhang, Chunyun, Weiyang Xu, Fanyu Meng, Hongyan Li, Tong Wu, and Lixin Xu. "The Information Extraction Systems of PRIS at Temporal Summarization Track." In TREC. (2013).
- [2] Marcu, Daniel. "Discourse trees are good indicators of importance in text." *Advances in automatic text summarization* 293 (1999): 123-136.
- [3] Filatova, Elena, and Eduard Hovy. "Assigning time-stamps to event-clauses." In *Proceedings of the ACL 2001 Workshop on Temporal and Spatial Information Processing*. (2001):1-8.

- [4] Mani, Inderjeet, and Mark T. Maybury. "Advances in automatic text summarization, vol. 293." Camb MA (1999).
- [5] James Pustejovsky, Jessica Littman, Robert Knippen, and Roser Sauri. "Temporal and Event Information in Natural Language Text, Language Resources and Evaluation", 2005, 39(2-3):123-164.
- [6] Pustejovsky, James, Robert Knippen, Jessica Littman, and Roser Sauri. "Temporal and Event Information in Natural Language Text." *Language Resources and Evaluation* 39, no. 2–3 (May 2005): 123–164. doi:10.1007/s10579-005-7882-7.
- [7] Färber, M., and Adam J. Finding Temporal Trends of Scientific Concepts. In *Proceedings of the 8<sup>th</sup> International Workshop on Bibliometric-enhanced Information Retrieval (BIR 2019) (CEUR Workshop Proceedings)*, Vol. 2345. (2019):132-139.
- [8] Grüninger, Michael, and Zhuojun Li. "The time ontology of Allen's interval algebra." In *24th International Symposium on Temporal Representation and Reasoning (TIME 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, (2017):1-16.
- [9] Mohd Pozi, Muhammad Syafiq, Adam Jatowt, and Yukiko Kawai. "Temporal Summarization of Scholarly Paper Collections by Semantic Change Estimation: Case Study of CORD-19 Dataset." *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020 (August 2020)*. doi:10.1145/3383583.3398563.
- [10] Pustejovsky, James, Robert Knippen, Jessica Littman, and Roser Sauri. "Temporal and Event Information in Natural Language Text." *Language Resources and Evaluation* 39, no. 2–3 (May 2005): 123–164. doi:10.1007/s10579-005-7882-7.
- [11] Guo, Qi, Fernando Diaz, and Elad Yom-Tov. "Updating Users About Time Critical Events." *Advances in Information Retrieval* (2013): 483–494. doi:10.1007/978-3-642-36973-5\_41.
- [12] Rautray, Rasmita, and Rakesh Chandra Balabantaray. "An Evolutionary Framework for Multi Document Summarization Using Cuckoo Search Approach: MDSCSA." *Applied Computing and Informatics* 14, no. 2 (July 2018): 134–144. doi:10.1016/j.aci.2017.05.003.
- [13] Alguliev, Rasim M., Ramiz M. Aliguliyev, Makrufa S. Hajirahimova, and Chingiz A. Mehdiyev. "MCMR: Maximum Coverage and Minimum Redundant Text Summarization Model." *Expert Systems with Applications* 38, no. 12 (November 2011): 14514–14522. doi:10.1016/j.eswa.2011.05.033.
- [14] Roser Sauri, Jessica Littman, Bob Knippen, Robert Gaizauskas, Andrea Setzer, and James Pustejovsky: *TimeML Annotation Guidelines: Version 1.2.1*, January 31, (2006).
- [15] Harabagiu, Sanda M., and Finley Lacatusu. "Generating single and multi-document summaries with gistexter." In *Document Understanding Conferences*, (2002): 40-45.
- [16] Sanampudi, Suresh Kumar, and G.Vijaya Kumari. "Temporal Reasoning in Natural Language Processing: A Survey." *International Journal of Computer Applications* 1, no. 4 (February 25, 2010): 68–72. doi:10.5120/100-209.
- [17] Guda, Vanitha, and SureshKumar Sanampudi. "Event Time Relationship in Natural Language Text." *International Journal of Recent Contributions from Engineering, Science & IT (IJES)* 7, no. 3 (September 25, 2019): 4. doi:10.3991/ijes.v7i3.10985.
- [18] Li, Wei, Xinyan Xiao, Jiachen Liu, Hua Wu, Haifeng Wang, and Junping Du. "Leveraging Graph to Improve Abstractive Multi-Document Summarization." *arXiv preprint arXiv:2005.10043* (2020).
- [19] Liu, Yang, and Mirella Lapata. "Hierarchical transformers for multi-document summarization." *arXiv preprint arXiv:1905.13164* (2019).
- [20] Yu N, Huang M, Shi Y, zhu x, Product re-view summarization by exploiting phrase proper-ties. In: *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, The COLING 2016 Organizing Committee, Osaka, Japan, (2016):1113–1124*.
- [21] Zhang, Yang, Yunqing Xia, Yi Liu, and Wenmin Wang. "Clustering Sentences with Density Peaks for Multi-Document Summarization." *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (2015). doi:10.3115/v1/n15-1136.
- [22] Ng, Jun-Ping, Praveen Bysani, Ziheng Lin, Min-Yen Kan, and Chew Lim Tan. "Exploiting Category-Specific Information for Multi-Document Summarization." In *COLING*, (2012): 2093-2108.