

Extractive Text Summarization for Social News using Hybrid Techniques in Opinion Mining

M. Nafees Muneera, P.Sriramya



Abstract: Presently almost all enterprises are oriented into building text data in abundance savoring the benefits of big data concept but the reality is that it's not practically possible to go through all this data/documents for decision making because of the time constraint. Here in exists intense need of an approach as an alternative for the actual content which can summarize the complete textual content. By adopting these summarizing approaches, the accuracy in data retrieval of summarized content via search queries can be enhanced compared to performing search over the broad range of original textual content. There are many text summarization techniques formulated having their own pros and cons. The present work focuses on a comprehensive news review of extractive text summarization process methods and also taking into account, data appended dynamically. The existing work recommends a technique of hybrid text summarization that's a blend of CRF (conditional random fields) and LSA (Latent Semantic Analysis) which being highly adhesive with low redundant summary and coherent and in-depth information. The above hybrid techniques is being extracted in five types that being: Positive and negative, statement, questions, suggestions and comments. The technique of LSA extracts hidden semantic structures within words/sentences that being commonly utilized in the process of summarization. The statistical modeling technique of CRF adopts ML (machine leaning) for offering structured detection and providing multiple options for evaluation of opinion summarization thereby identifying the most appropriate algorithm for news text summarizations considering the heavy volume of datasets.

Index Terms: Text summarization, conditional random fields, Latent Semantic Analysis, machine leaning, Feature selection, Hybrid text summarization technique, extractive text summarization.

I. INTRODUCTION

Today, the Internet holds magnificent data volume and retrieving significant information online sees almost

impractical. Nowadays, online news is available that's being offered by various online newspapers, social media applications and public authority websites. The content published by them includes social happenings, economy, politics or covering debatable topics that prevail amidst general public. As the news is composed of relatively huge set of data, summarizing such a massive content becomes an impractical and tedious job. As a way out, technique of Text summarization can be adopted that works by formulating a condensed form of the original document automatically and simultaneously keeping intact the original document in a short format resembling throughout meaning. Text summarization can be achieved using the methods of abstractive and extractive summarization. The abstractive summarization method adopts techniques of natural language generation that generates summary illustrating the internal semantics. On the other hand the method of extractive summarization generates summary by cohesively gathering similar sentences from the text documents. There are many text summarization techniques formulated having their own pros and cons. The extractive text summarization ignores the issues faced in the past works such as identifying extremely significant information chunks within the news reviews, eliminating unwanted information and allowing the user to view the article's positive and negative side and reducing the process details. The research work targets on the extractive summarization of news reviews and determining of sentiments on news article based on users query, such as Positive and negative statements, suggestions, comments, questions alongside taking into account data appended dynamically. The existing research recommends a technique of hybrid text summarization that's a blend of CRF (conditional random fields) and LSA (Latent Semantic Analysis) for extracting review types such as Positive and negative statements, suggestions, comments and questions. The statistical modeling technique of CRF adopts ML (machine leaning) for offering structured detection of user reviews. It's employed for parsing sequential data and for encoding familiar relationships amidst observations. The technique of LSA extracts hidden semantic structures within words/sentences. The system handling processes includes:

- Extractive Summarization
- Collection of datasets
- Pre-processing, dynamic adding data
- Feature selection
- Hybrid text summarization technique
- CRF (Conditional random field) and
- LSA (Latent Semantic analysis).

Revised Manuscript Received on February 05, 2020.

* Correspondence Author

Nafees Muneera*, Dept of Computer Science and Engineering, Saveetha School Of Engineering ,Saveetha Institute Of Medical and Technical Sciences, Chennai, India, .

Sriramya, Computer Science and Engineering, Saveetha School Of Engineering ,Saveetha Institute Of Medical and Technical Sciences, Chennai, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

On comparing the proposed technique of hybrid text summarization with other existing techniques, better results are generated with the proposed one. Following is the research structure: Section 2 describes related work. Section 3 presents the proposed technique. Section 4, illustrates results and further discussions. At last the research concludes with Section 5.

II. RELATED WORK

Sagar M. Patel [1] illustrates the technique of extractive based summarization. For the purpose of extractive based automatic text summarization, the approach of lexical chain analysis is proposed by the author that includes sample generated lexical chains, so that the content's summary can be obtained. Moreover linguistic pre-processing steps are carried out by making use of existing tools and plugins. According to the output generated, the lexical chain approach yields in improvised outcome compared to others. Though it faces the drawback of duplication in summary sentences.

Santosh Kumar Bharti [2] recommends a hybrid technique for automatically extracting keyword concerning multi-document text summarization related to e-newspaper articles. Describing or manually extracting relevant keywords can be quiet tedious. Using the proposed approach of Automatic keyword extraction, words and phrases can be chosen from the text document which depicts the underlying sentiment of the document. The proposed automatic keyword extraction technique is compared with TF-IDF, TF-AIDF and NFA keyword extraction techniques. Though the proposed system offers the benefit of lower cost but has the drawback of being complicated and lacking semantic analysis of source text.

Yong Zhang [3] says over the approach of multi-document extractive summarization. Earlier methods rely upon 1.) Hand-crafted features for representing sentences; 2) capturing syntactic and semantic information simultaneously within documents is not feasible. 3) Ability to summarize is poor. In this work, the author resolves the issue of framing difficult hand-crafted features and improving summarization skills. CNNs (convolutional neural networks) or multi view CNNs are being built which aids in jointly ranking the sentences and also acquire their features. Multi view CNN is basically adopted for higher and improvised learning of original convolutional neural networks.

Kim Schouten [4] surveys Aspect-Level Sentiment for identifying and combining sentiments on articles highlighted in the documents or their aspects. Also detailed review of the existing state-of-the-art is being provided. Various techniques illustrated in the survey involves ML (machine learning) approaches such as Frequency-Based Methods, supervised and unsupervised learning and syntax based model. According to the output and discussion the aspect-level sentiment (which being concept-centric and semantically-rich) is remarked as the best assuring future research approach.

Oi-Mean Foong [5] illustrates text summarization by employing LSA (latent semantic analysis) model in android based mobiles. Text summarization process can be facilitated automatically by employing numerous available tools but the hitch being that semantic summary is not achieved. An automatic text summarization is proposed by the author by adopting LSA (latent semantic analysis) which results in short summary pertaining to every document. The system

yields in increased performance which being an advantage but also needs extra storage space which is a disadvantage. According to the results, the summary output generated by LSA has an average F-score of 0.386.

Raj Kumar V.S [6] presents a survey on text summarization by employing optimization algorithm. Earlier work suffers from generating inefficient summary and information-organization. Various optimization algorithms being proposed includes: artificial bee colony algorithms, particle swarm optimization, colony optimization and genetic algorithms for producing better/improved summary and information-organization. According to the output, particle swarm optimization assures to be highly effective and improves the functionality of text summarization.

Shankari V. Gajul [7] performs discussion on the Text Summarization related to News Articles by employing Modified NN (Neural Networks). Usually, the news articles lacks summaries associated with them also the titles are inadequate for expressing the important aspects. A ML approach is proposed which employs ANN (artificial neural networks) for generating summaries of news articles with random length. Techniques being incorporated are NN (neural network) training, sentence selection and feature fusion. According to the results, the performance of the proposed system is better over the test paragraphs, yielding an accuracy of around 90 %.

Shail Shah [8] proposes an Query Based Text Summarizer. Earlier approaches were depended upon linguistic or statistical techniques. Here, a graph structure is proposed to be utilized in the model to naturally convert the meaning and structure related to a cohesive text within the document. TF-IDF helps in performing the scoring of sentence and keyword occurring in the query and in order to rank the sentences based on their importance, Text Rank algorithm is adopted. At last the total output produces the abstractive summary relying upon the query. High accuracy is obtained using the above approach.

Mondher Bouazizi [9] recommends a conventional multi-class classification, which is being executed on data set gathered from social networking site, Twitter. All the sentiments associated with an online post are identified that is, Online Social Media content, Multi-class sentiment analysis and "quantification". The technique automatically assigns distinctive scores to every sentiment in a tweet. Thereafter sentiments having the maximum score is chosen which is analysed and conveyed in the text. To achieve this, SENTA tool is being utilized with essential elements for executing the above process. The outcome reveals that the task is feasible with an F1 score of 45.9%.

Shrabanti Mandal [10] proposed the approach of Hybrid text summarization on the basis of extractive summarization and adopts the approach of statistical measure, sentiment analysis and eventually incorporates fuzzy logic technique for the selection of sentence. The summary is produced depending upon the importance level of the sentence. According to the output, hybrid approach is better in performance compared to rest others. It offers the benefit of 'search of particular information', projecting users interest. The shortcoming being that there is absence of detailed information.

Kuan-Yu Chen [11] discusses Extractive Broadcast News Summarization Leveraging Recurrent NN (Neural Network). Prime issues confronted by Language Modeling approach is building of the sentence models and precisely estimating every sentence's parameters within the document for summarization. The author proposes an original framework of RNNLM (recurrent neural network language modeling) for the purpose of extractive broadcast news summarization. Assumed sentence models reduces word usage indications along with long-span structural information of word co-occurrence associations concerning broadcasting news documents.

Kamal Al-Sabah [12] discusses regarding Hierarchical Structured Self-Attentive Model for HSSAS (Extractive Document Summarization). Conventional approaches of Text Summarization confronted various challenges like memory issues and adopting knowledge related to document structure. The above mentioned model of hierarchical structured self-attention mechanism aims to build up the sentences and document embeddings. The model works by mirroring the document's hierarchical structure thereby helping to acquire enhanced representation of features. Popular datasets were utilized namely, CNN / Daily Mail, and DUC 2002 for evaluation of the proposed model.

K.Aruna [13] examines sentiment analysis referring to information available on social media by adopting DM (data mining) techniques. Social networking sites are being tracked by the author for extracting users' opinion on varied topics, for evaluating DM (data mining) techniques adopted in analysis of social network and also to get an overall review of present and past researches being conducted on sentiment analysis across OSN sites like Facebook, twitter etc... The output generated reveals the use of DM techniques in the analysis of social network.

Xinjie Zhou [14] discusses the approach of Opinion Extraction and Summarization concerning Chinese Micro blogs. The proposed opinion extraction system is referred to as CMiner which basically attends to complex opinion extraction/mining tasks, opinion target extraction and opinion summarization. For the opinion target extraction, algorithm of unsupervised label propagation is adopted and a dynamic programming based algorithm is adopted for opinion summarization. For producing representative targets as well as summary sentences related to each group, the algorithm of co-ranking is incorporated.

Dandan Jiang [15], proposes a novel approach for carrying out sentiment computing for news events and WEAN (Word Emotion Association Network) is being formed for conveying together their semantics and emotions. The prevailing sentiment computing methods usually relies upon the supervised methods or standard emotion thesaurus, which being non-scalable. Using the proposed approach of WEAN, initial word emotions can be acquired which is filtered using the standard emotion thesaurus.

III. PROPOSED WORK

A. Overview

The process of web mining involves adopting DM (data mining) techniques for straight away mining/extracting information from the Web. Initially the news articles dataset is assembled online via website. Thereafter, reviews related to news article are fetched from it. Then comes the process of

pre-processing which eliminates unnecessary data from the reviews while allowing dynamic appending of news data. Next is the process of feature selection which helps in choosing significant data while avoiding the unnecessary ones. Then is the application of hybrid text summarization that's a blend of CRF (conditional random fields) and LSA (Latent Semantic Analysis) that offers a structured detection for the identification of relationship amidst texts and label it. The technique of LSA extracts hidden semantic structures within words/sentences. It delivers summarization of news reviews and determining of sentiments on news article based on users query, such as Positive and negative statements, suggestions, comments, questions.

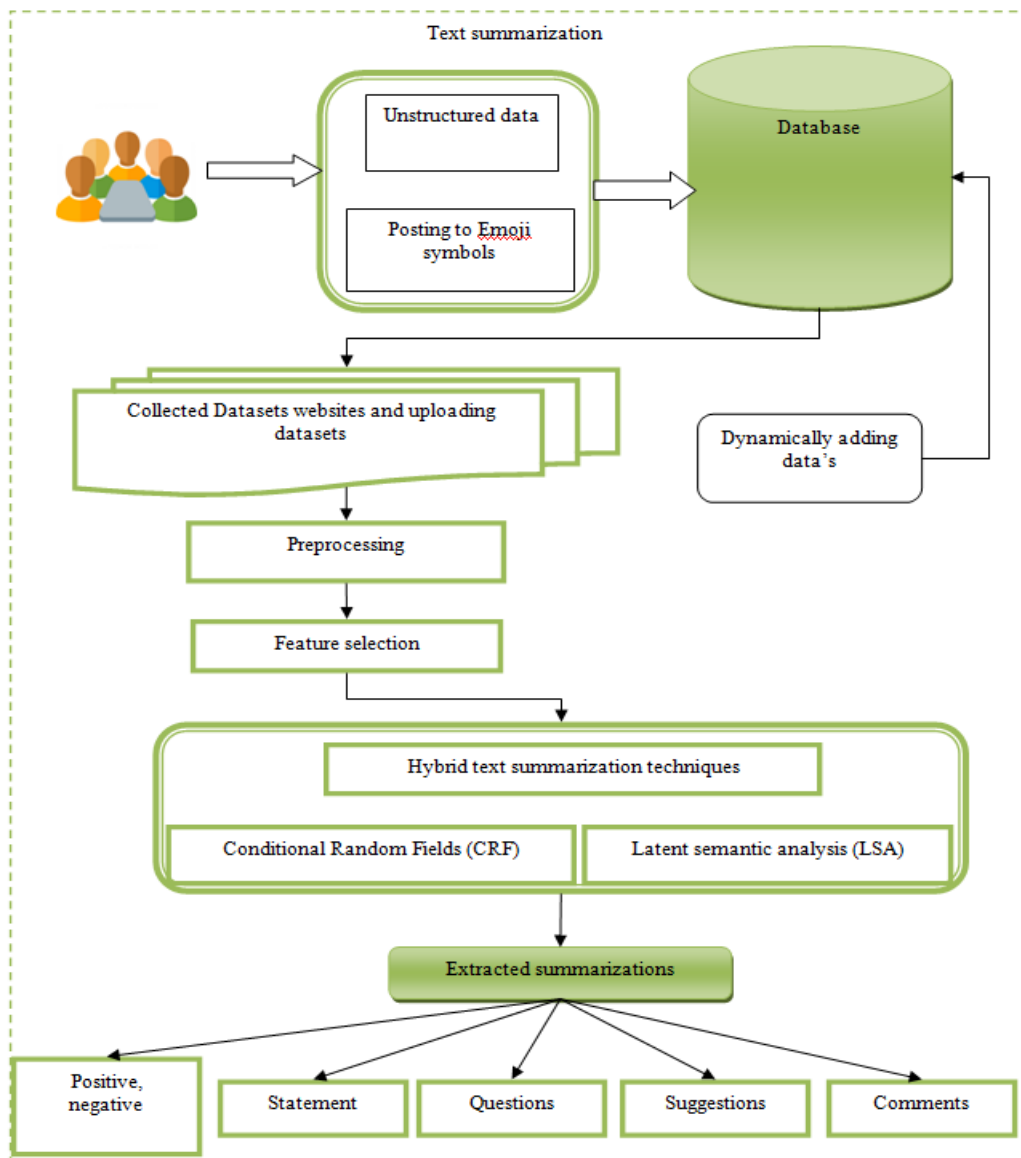


Figure 1: Overall proposed Architecture

B. Extractive Summarization

Extractive Text Summarization methods involve methods of Unsupervised Learning and Supervised learning. It withholds the summary of significant sentences concerning a particular article. The summary content and the collection of sentences depend on reader’s need and demand. Based on the importance, the sentences are directly selected from the original document.

C. Collection of Datasets

The News article is considered as a piece of writing to be published within a newspaper or magazine. Large quantity of news reviews is being gathered online via website that holds data in structured, un-structured and semi-structured form. News review basically involves analysis concerning specific News article and are represented as texts and Emoji. Emoji is refers to as smileys and ideograms that are being made use of in sending e- messages.

D. Pre-Processing

Once data is assembled it undergoes pre-processing. There is presence of unnecessary and crap data in the collected news reviews which causes hindrance in the process of evaluation. Pre-processing helps in eliminating this disturbing and insignificant data from the news reviews. Initially the data cleaning takes place to manage missing data, identifying and deleting repeated data, thereafter data scrubbing helps in eliminating noise and irrelevant data present in the reviews.

E. Dynamic adding data

Users reviews and opinions concerning news articles is in massive numbers which keeps on rising. All these news related reviews are stored within database which actually holds important details regarding the news article. Also there is dynamic addition of reviews on a daily basis which aids in better analysis of news article and to obtain precise review of the concerned news article. The stored data is thereafter useful in the process of feature selection.



F. Feature Selection

Feature selection attempts to pick out essential features which can minimize data dimensionality and also enhance the performance. It strictly chooses the most appropriate and significant words and on the other hand get rid of the duplicate, unnecessary and insignificant features as they don't offer any valuable information compared to the selected features. Hence by using Feature selection, the most suitable set of features can be enlisted.

G. Hybrid Text Summarization Techniques

This technique helps in identifying the structural format (which is grammatical) of the reviews, identifying then underlying meaning among the words present in the reviews and determining reviews sentiments whether positive and negative statements, suggestions, comments and questions. There exist numerous techniques of Text summarization which basically generates the content's summary. The existing work recommends a technique of hybrid text summarization that's a blend of CRF (conditional random fields) and LSA (Latent Semantic Analysis). The technique of CRF adopts ML (machine learning) for offering structured detection of user reviews and the technique of LSA extracts hidden semantic structures within words/sentences of the document.

H. Conditional Random Fields

CRFs (Conditional Random Fields) can be considered as a discriminative undirected probabilistic graphical model which is adopted for parsing/labeling the sequential data. It's a subset of the sequence modeling family. CRF helps in predicting the labels sequence for sequencing the input samples. Every word/sentence is being labeled by CRF depicting what it means. But before it performs the labeling of sentences, it must first determine a set of feature functions. Conditional Random Fields framework involves two stages first is the parameter estimation and second being the feature consideration. The first stage of Parameter Estimation implies estimating a set of weights by employing the process of maximum likelihood. The factors considered by the Feature function involves: input sentence - s, position of a word in the given sentence, the label (L_i) of the existing word, the label (L_{i-1}) of the prior word. Thereafter, each feature function (f_j) is allocated a weight - λ_j. For the given sentence s, it's possible to score a labeling l of s by totaling the weighted features for complete set of words in the sentence - λ:

$$score(l|s) = \sum_{j=1}^m \sum_{i=1}^n \lambda_j f_j(s, i, l_i, l_{i-1})$$

The initial sum is applied over each feature function - j, and thereafter the inner sum applies over every position - i of the sentence. At last the scores obtained can be converted in form of probabilities p(l|s) ranging between 0 and 1 by normalizing and exponentiation:

$$p(l|s) = \frac{\exp [score(l|s)]}{\sum_{l'} \exp [score(l'|s)]} = \frac{\exp [\sum_{j=1}^m \sum_{i=1}^n \lambda_j f_j(s, i, l_i, l_{i-1})]}{\sum_{j=1}^m \sum_{i=1}^n \lambda_j f_j(s, i, l'_i, l'_{i-1})}$$

- f₁(s,i,l_i,l_{i-1}) = 1, if l_i = good having Thumbsups (i.e., positive sentences) ; else 0. ** If the weight (λ₁) linked

to the feature is big and positive, then the feature signifies that, labelling preferred in case of words having good and emoji , thumbsup is labelled as Positive.

- f₂(s,i,l_i,l_{i-1}) = 1, if i=1l_i= Bad, and having thumbsdown (i.e., negative sentences); else 0. ** here, if the weight (λ₂) linked to the feature is big and negative, then labelling preferred in case of words having bad and thumbsdown emoji is labelled as negative.
- f₃(s,i,l_i,l_{i-1})=1 if l_{i-1}= opinions and l_i=ideas; **weights(λ₃) for this label holds comments and suggestions.
- f₄(s,i,l_i,l_{i-1})=1 if l_{i-1}= query and l_i=expressions. **Weight (λ₄)for this label holds Questions and Statement.

I. Latent Semantic Analysis

The technique of LSA (Latent Semantic Analysis) extracts hidden semantic structures within words/sentences which being commonly utilized in the process of text summarization. It's an approach of unsupervised learning which need not require explicit training or knowledge. The Latent Semantic Analysis basically converts the original data and forms new variables in order to minimize the dimensionality, which helps in achieving increased computing speed. There exist numerous set of words in the news article reviews which can depict similar or different meaning. There are situations where one word can exhibit different meanings and there can be different words having same meaning. So LSA (Latent Semantic Analysis) helps in determining meaning of various words that exists in the news review. It takes into account the input text of the document along with extracted details like frequently occurring joint words and words appearing often in multiple sentences. A large quantity of common words within the sentences depicts that the sentences are linked semantically. SVD (Singular Value Decomposition) is adopted to identify the relationship among the words and sentences, helping in minimizing the noise thereby enhancing the accuracy and performance. Singular Value Decomposition is a mathematical matrix decomposition techniques which help in detecting relationships patterns among the words and sentences of the text document and also identifying similarity in the meanings associated with words/sentences. LSA is divided into 3 major levels.

These steps are as follows:

i. Forming of input matrix:

Text/input reviews is denoted as a matrix. Each row depicts a word and the column depicts the sentence. The cell value signifies the word's importance.

ii. Singular Value Decomposition (SVD):

Singular value decomposition (SVD) is a mathematical matrix decomposition method implemented on the input matrix and is used to identify patterns in the relationships between the terms and sentences.

In terms of mathematical equation the SVD can be denoted as an m×n matrix (M). M is represented as



$$M = U \Sigma V^T$$

Here U denotes an m×n matrix signifying the original rows as vectors of extracted values

Σ represents an n×n rectangular diagonal matrix having nonnegative real numbers on the diagonal depicting the scaling values and VT (the conjugate transpose of V) denotes an n×n real or complex unitary matrix which signifies the original columns as vectors of extracted values.

iii. **Sentence Selection:**

Once SVD is performed, the output generated is utilized to choose the sentences.

J. Category Summarization

Extractive helps in picking a group of significant sentences within the original document, thus concatenating the same to formulate a summary. The technique of hybrid text summarization that’s a blend of CRF (conditional random fields) and LSA (Latent Semantic Analysis) for extracting review types such as Positive and negative statements, suggestions, comments and questions for determining news article review of prior users. Depending on the users query, the appropriate news article is fetched. Just by viewing the news, the most necessary sentiments can be determined by the user without actually going through the entire news.

K. Advantages

1. The news article can be easily identified.
2. Minimizing the user’s access time in examining the article.
3. Depicts various sentiments that includes : Positive, Negative, Question, Statement Comments.
4. Technique of hybrid summarization yields in high accuracy.

IV. RESULT AND DISCUSSION

Technique of Text summarization works by formulating a condensed form of the original document automatically and simultaneously keeping intact the original document in a short format resembling the overall meaning. With the existence of endless information over the internet, finding the appropriate desired information amidst a large variety tends to be complicated and impractical also the retrieved information consists of insignificant and junk data. Also it exhibits user’s sentiments such as positive/negative sides of the articles which makes the article reading unfeasible for the user. The research proposes the technique of hybrid text summarization that’s a blend of CRF (conditional random fields) and LSA (Latent Semantic Analysis).CRF provides structured detection of user reviews. The technique of LSA extracts hidden semantic structures within words/sentences. The hybrid text summarization is adoptedfor extracting review types such as Positive and negative statements, suggestions, comments and questions. The proposed technique is compared with rest existing approaches and the hybrid approach yields promising results with enhanced output.

.Table 1: Comparison of Text Summarization Techniques

S.NO	Techniques	Accuracy (%)	Time (s)
1	Kernel SVM	76.1	4.2
2	ANN	80.3	3.5
3	Multiclass Liner Kernel	87.4	2.4
4	Hybrid Text Summarization Technique	92.4	1.1

Table 1 depicts the comparison and overall performance of Text summarization Techniques with rest of the techniques namely, ANN, Kernel SVM, Multiclass Liner Kernel with Hybrid summarization. The hybrid approach yields promising results with enhanced output. Also, other factors like accuracy and time are compared.

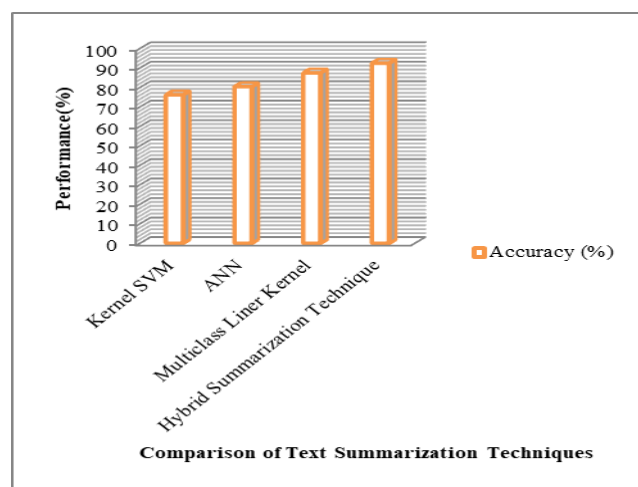


Figure 2: Comparison of Text Summarization Techniques

Fig 2, depicts the comparison and overall performance of Text summarization Techniques with rest of the techniques namely, ANN, Kernel SVM, Multiclass Liner Kernel with Hybrid summarization. The hybrid approach yields promising results with enhanced output. Also, other factors like accuracy and time are compared.

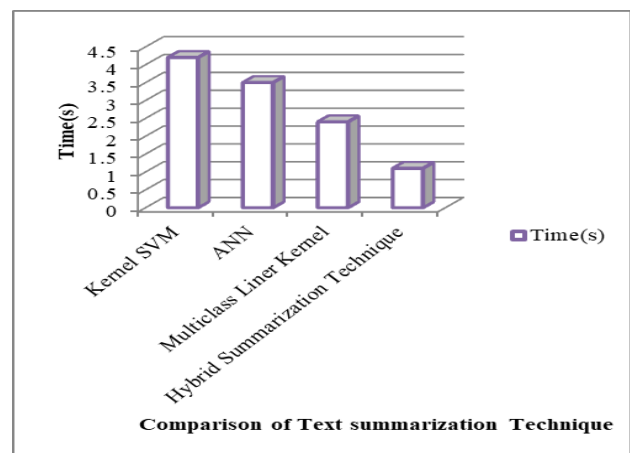


Figure 3: Comparison of time with existing

V. CONCLUSION

Extractive text summarization generally utilizes features generated from the text content like 'frequency'. Various challenges confronted by Extractive summarization includes retrieval valuable chunk of information from the article, eliminating insignificant information, reducing process details and that only the positive and negatives article list is accessible to the user. The proposed technique of hybrid text summarization is adopted for extracting review types such as positive and negative statements, suggestions, comments and questions. The proposed approach yields in an accuracy of 92.4% which being higher than rest of the available techniques. Also the technique assures to be helpful to the user. According to the output generated the proposed approach generates effective and overall improvised results.

REFERENCES

1. Sagar M. Patel, Vipul K. Dabhi, Harshadkumar B. Prajapat, "Extractive Based Automatic Text Summarization", Journal of Computers, Vol. 12, No. 6, pp. 550-563, 2017.
2. Santosh Kumar Bharti, Korra Sathya Babu, Anima Pradhan, "Automatic Keyword Extraction for Text Summarization in Multi-document e-Newspapers Articles", European Journal of Advances in Engineering and Technology, pp. 410-427, 2017.
3. Yong Zhang, Meng Joo, Rui Zhao, Mahardhika Pratama, "Multiview Convolutional Neural Networks for Multi-document Extractive Summarization", IEEE Transactions On Cybernetics, pp. 1-13, 2016.
4. Kim Schouten, Flavius Frasinca, "Survey on Aspect-Level Sentiment Analysis", IEEE Transactions on Knowledge and Data Engineering, pp. 1-20, 2015.
5. Oi-Mean Foong, Suet-Peng Yong, Farha-Am Jaid, "Text Summarization Using Latent Semantic Analysis Model in Mobile Android Platform", IEEE Computer Society, pp. 35-39, 2015.
6. Raj Kumar V.S, Chandrakala. D, "A Survey On Text Summarization Using Optimization Algorithm", Elk Asia Pacific Journal Of Computer Science And Information Systems, Vol. 2, Issue. 1, 2016.
7. Shankari V. Gajul, "A Survey on Text Summarization for News Articles Using Modified Neural Networks", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 6, Issue 7, pp. 13098-13105, 2017.
8. Shail Shah S, Adarshan Naiynar B, Amutha, "Query based text summarization", ARPN journal of engineering and applied sciences, Vol. 12, pp. 5456-5468, 2017.
9. Mondher Bouazizi, Tomoaki Ohtsuki, "Multi-Class Sentiment Analysis in Twitter: What if Classification is not the Answer", IEEE, pp.1-16, 2018.
10. Shrabanti Mandal, Girish Kumar Singh, Anita Pal, "A Hybrid Text Summarization Approach", Journal of Informatics and Mathematical Sciences, Vol. 9, No. 3, pp. 547-555, 2017.
11. Kuan-Yu Chen, Shih-Hung Liu, Berlin Chen, Hsin-Min Wang, Ea-Ee Jan, Wen-Lian Hsu, Hsin-Hsi Chen, "Extractive Broadcast News Summarization Leveraging Recurrent Neural Network Language Modeling Techniques", IEEE/ACM transactions on audio, speech, and language processing, Vol. 23, pp. 1322-1334, 2015.
12. Kamal Al-Sabahi, Zhang Zuping, Mohammed Nadher, "A Hierarchical Structured Self-Attentive Model for Extractive Document Summarization (HSSAS)", IEEE, pp. 1-9, 2018.
13. K. Aruna, S. Anupriya, "Sentiment Analysis On Social media Information Using Data mining Techniques A Review", International Journal of Pure and Applied Mathematics, Vol. 120, No. 6, pp. 1-9, 2018.
14. Xinjie Zhou, Xiaojun Wan, and Jianguo Xiao, "CMiner: Opinion Extraction and Summarization for Chinese Microblogs", IEEE Transactions on Knowledge and Data Engineering, pp. 1-14, 2016.
15. Dandan Jiang, Xiangfeng Luo, Junyu Xuan, Zheng Xu, "Sentiment Computing for the News Event Based on the Social Media Big Data", IEEE, pp.1-10, 2016.

AUTHORS PROFILE



M.Nafees Muneera, Research Scholar, Dept. of Computer Science and Engineering, Saveetha School Of Engineering, Saveetha Institute Of Medical and Technical Sciences, Chennai, India.. Her research interest is in data mining and data analytics.



Dr. P.Sriramya, Associate Professor in the Department of Computer Science and Engineering has more than 6 years of teaching experience and 5.5 years of Industrial experience.. Her research interests include Algorithms, Optimization Techniques, Network Security and Artificial Intelligence and Neural networks. She received Young Women Educator & Scholar Award from NFED and Young Women Achiever Award from Venus International Foundation She has 16 publications which are all in Scopus indexed and listed in Anna University Annexure.