

# Automatic Pre-Processing of Marathi Text for Summarization



Apurva D. Dhawale, Sonali B. Kulkarni, Vaishali M. Kumbhakarna

**Abstract:** The text summarization is a technique where the original large text is condensed into smaller version without changing its abstract meaning. The text summarization is done on the common foreign and regional languages typically, but infrequent work has been observed for the Marathi language. As the amount of e-contents on web is increasing drastically, the users are facing difficulty to read the newspaper articles with extraction of its different perspectives with sorting. We are focussing on educational, Political and sports news for summarization, which will be helpful for students who are appearing for competitive exams. This paper explores the pre-processing techniques for Marathi e-news articles.

**Keywords:** Text summarization, POS tagging, Pre-processing, LDA(Latent Dirichlet Allocation), LNS (Label Induction Grouping), SVM (Support Vector Machine)

## I. INTRODUCTION

Summarization is defined as the extraction of features of text document and generating abstract with same meaning. [1] To have an access to reliable and accurate data, user needs to implement a very potent system which will give best results. The summarization of text is an interesting area where people of 21<sup>st</sup> century would be relying for time saving, accuracy, & reduced efforts for reading the whole document. There are many prominent languages on which the work has been done in the area of text summarization. But today the need for regional language text summarization is very much obligatory. Keeping this in mind, the work for regional languages in Maharashtra has been reviewed, where the Marathi Language is a bit less focussed. The literature for Marathi Language text summarization shows that there is no observed powerful tool, or system which gives high efficiency in summarizing Marathi text. Soit's needed to focus on the Marathi language text summarization. There are two major steps through which the text goes for the efficient output, a) Pre-processing&b) Processing.<sup>[3]</sup>

## II. LITERATURE STUDY

To find appropriate information, a user needs to search through the entire documents this causes information overload problem which leads to wastage of time and efforts, and this happens when user queries for information on the internet he may get thousands of result documents which may not necessarily relevant to his concern.

Revised Manuscript Received on October 10, 2020.

\* Correspondence Author

**Ms. Apurva D. Dhawale\***, Department of Computer Science, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India.

**Dr. Sonali B. Kulkarni**, Completed her Master of Science, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India

**Ms. Vaishali M. Kumbhakarna**, Completed Master of Science, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India

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

Retrieval Number: 100.1/ijeat.A18031010120

DOI:10.35940/ijeat.A1803.1010120

Journal Website: [www.ijeat.org](http://www.ijeat.org)

To deal with this dilemma, automatic text summarization plays a vital role. Automatic summarization condenses a source document into meaningful content which reflects main thought in the document without altering information [13]. There are distinctive automatic text summarization systems existing for most of the regularly used natural languages. [4] The Text summarization methods can be categorized by the way it is done. The approaches mainly include single document, multi document, monolingual, multi lingual, generic, query based, indicative, informative summary.[14] These methods are used for numerous foreign and Indian languages all over world. As we are focussing on Marathi language, which is the regional language of Maharashtra the following work has been done in recent years: Mr. Shubham Bhosale, Ms. Diksha Joshi, Ms. VrushaliBhise, Prof.Rushali A. Deshmukh [1] proposed a system for Marathi newspaper text summarization using Ranking algorithm which gives average of 30% to 40 % size of original article. Anishka Chaudhari<sup>1</sup>, Akash Dole<sup>2</sup>, Deepali Kadam, proposed a system which translates Marathi dataset to English using Google Translate API and then summarizes news articles using a bi-directional encoder-decoder LSTM model. The resultant summary is again translated to Marathi using Google Translate API.[5] Pooja Bolaj,SharvariGovilkar[2] developed a text classification system for Marathi documents using supervised learning methods & ontology based classification technique which classifies Marathi documents belonging to Festival class i.e. Diwali. Deepali K. Gaikwad, Deepali Sawane and C. Namrata Mahender, developed a system for rule Based Question Generation for Marathi Text Summarization using Rule Based Stemmer. The paper shows technique which is used for generation of the appropriate question on given input/text.[6] Yogeshwari V. Rathod [7] used sentence ranking algorithm to generate summary of Marathi news articles by extractive method. It gives effective summary in less time and with least redundancy. Shradha A. Narhari, RajashreeShedge [8] proposed a text categorization of Marathi documents using LINGO & PCA algorithm. They proved this with improved results. Jaydeep Jalindar Patil, Prof. NagarajuBogiri[9] used LINGO [Label Induction Grouping] algorithm for improving results efficiently in Marathi text documents. Prakhar Sethi, Sameer Sonawane, SaumitraKhanwalker, R. B. Keskar [10] developed a system to Overcome the limitations of the lexical chain approach to generate a good summary using the WordNet thesaurus, pronoun resolution for news articles. N. Dangre, A. Bodke, A. Date, S. Rungta, S.S. Pathak [11] proposed a System for Marathi News Clustering using Cluster algorithm to collect relevant Marathi news from multiple sources on web which



results in enabling rich exploration of Marathi contents on web. Mamatha Balipa, Dr. Balasubramani R, Harolin Vaz, Christina Shilpa Jathanna, attempted summarizing information from online health care forums about the disease Psoriasis to implement automatic text summarization. Online text is extracted using BeautifulSoup class available in urllib2 module.

Then the topic of the text is confirmed to be Psoriasis by using Latent Dirichlet Allocation (LDA) algorithm.[20]

Chirantana Mallick, Ajit Kumar Das, Madhurima Dutta, Asit Kumar Das and Apurba Sarkar, proposed a method which constructs a graph with sentences as the nodes and similarity between two sentences as the weight of the edge between them.[21] Reda Elbarougy, Gamal Behery, Akram El Khatib, applied modified page rank algorithm with an initial score for each node that is the number of nouns in this sentence. More nouns in the sentence mean more information, so nouns count used here as initial rank for the sentence. Edges between sentences are the cosine similarity between the sentences, to get a final summary that contains sentences with more information and well connected with each other. [22] Ahmed Elrefaiy, Ahmed Rafat Abas, Ibrahim Elhenawy, provided a review of collaborative survey which focuses on unsupervised techniques. It also describes evaluation of techniques of the summaries.[23]

Rasim Alguliev, Ramiz Aliguliyev, shown an approach which can improve the performance compared to state-of-the-art summarization approaches. They have proposed new criterion functions for sentence clustering. They also have developed modified discrete differential evolution algorithm to optimize the objective functions.[24] Kalliath Abdul Rasheed Issam, Shivam Patel, Subalalitha C. N., proposed technique which aims to capture all the varied information present in source documents. Also they have discovered that their model produces encouraging ROUGE results and summaries when compared to the other published extractive and abstractive text summarization models. [25] Siddhant Upasani, Noorul Amin, Sahil Damania, Ayush Jadhav, A. M. Jagtap, obtained the rank or score of each sentence and the sentences with the rank above a particular value can be chosen to be included in the summary.[26] Yash Asawa, Vignesh Balaji, Ishan Isaac Dey, surveyed numerous approaches, merits and limitations of the techniques of summarization. The Benchmark datasets of this domain and their features have also been examined. [27]

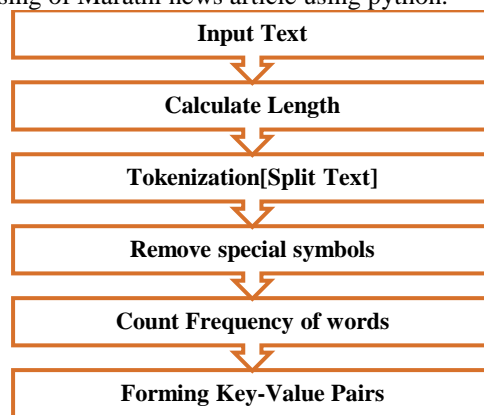
### III. PROPOSED SYSTEM

There are multiple types of text summarization which includes bilingual, multilingual, single document, multi document text summarization where the categories can be: 1] Foreign Language & 2] Indian language. Literature survey in the paper shows that the Foreign language text summarization is done using sentence ranking, deep learning, word frequency and distribution, fuzzy inference system, rule based, Genetic algorithm, LDA (Latent Dirichlet Allocation), Random Indexing and page rank algorithms. Indian Language text summarization is done using Scoring of sentences, ROUGE evaluation toolkit, Sub graph, Language-Neutral Syntax (LNS), Support Vector Machine (SVM) classifier, hybrid algorithm, Bernoulli Model of Randomness algorithms. [12] Here we are focussing on the Marathi text processing which can be done by using several algorithms which are Text ranking, LINGO,

Supervised Learning Method, Clustering, lexical chain, domain specific summarization algorithms.[12] Sheetal Shimpikar, Sharvari Govilkar, worked on approach which takes Marathi documents as input text. The first step is pre-processing of the input text & used rich semantic graph method. They proved that the Rich Semantic Graph based method gives the correct, bug free result.[16]

In a nation like India there are 22 languages spoken, which are written in 13 different scripts, with about 720 dialects. Taking this into consideration developing a nationwide summarization tool for India would be a very difficult problem. Jovi D'silva, Dr.Uzzal Sharma examined approaches to this problem and also highlight some existing research that has been done in Indian languages. They proved a language independent approach for text summarization can prove to be enormously constructive as the algorithm would have the potential to create summaries irrespective of the language of the input text.[17] Poonam Kolhe, Prof. Ashish Kumbhare, designed an algorithm that can recognize the action word by abstraction and summarize the input document by extraction and attempting to modify this extraction using a NLP tools like WordNet.[18] Umakant Dakulge, S. C. Dharmadhikari, proposed a framework which summarizes a single document using extraction method. The TF-ISF, sentence length, sentence positional value, SOV verification are used to make the summary more relevant and precise. [19] In this research, we are using extractive based approach using Text ranking algorithm where the document is read first, its length is calculated, and it would generate a summary which gives us important sentences according to the requirement of the user. The relevant literature shows that there are many methods & algorithms suitable for Text processing and text summarization as the digital text is gaining importance day by day. The result may vary depending on the language chosen and the selected algorithm.

Marathi is considered as an Indo-Aryan language. The people of Maharashtra speak this language primarily. Marathi is morphologically rich so the classification of text gets very difficult. [2] The steps below show the pre-processing of Marathi news article using python.



**Fig.1. Pre-processing of Marathi news article**

#### A. INPUT TEXT

The first step for text processing is input the text or paragraph for summarization. The input text may contain words,



sentences or paragraphs. The validity of text is checked and if there are some words or sentences which are not in Marathi language, they are eliminated from the document and then it is sent for further processing.

```
mytext= "" "केद्रीयमाध्यमिकशिक्षणमंडळ'तर्फे (सीबीएसई)
इयत्तादहावीचानिकालआज,
१५जुलैलाजाहीरकेलाजाणारआहे.निकालाचीनेमकीवेळमात्रबो
र्डनिजाहीरकेलेलीनाही. बारावीचानिकालसोमवारी,
१३जुलैलाजाहीरझालाहोता.दहावीचानिकाल
'सीबीएसई'चेअधिकृतसंकेतस्थळ cbse.nic.in येथेदिसेल.
यासोबतचनिकालासाठीस्वतंत्रपेजअसलेल्या
cbseresults.nic.in यालिंकवरहीतोपाहतायेईल.
यंदा
'सीबीएसई'नेआयव्हीआरएससुविधाउपलब्धकरूनदिली
आहे. यासाठीविद्यार्थ्यांना०११-२४३००६९९, ०११-
२८१२७०३०याक्रमांकांवरसंपर्कसाधावालागेल.कॉलसुरू
असतानाचविचारलेगेल्यावरमोबाइलवरआपलारोलनंबर
आणिजन्मतारीखटाकल्यानंतरनिकालसमजेल. ""
```

## B. PRE-PROCESSING

In Natural Language Processing(NLP), one of the important and traditional step is to pre-process the input text. It transforms the text in more comprehensible form by which the machine learning algorithms work well with text. Basically, the unstructured data is turned into structured one. If we do not apply pre-processing then data would be very inconsistent and could not generate good analytics results.[15] Here we are installing Python Libraries which work with NLP & Information retrieval for our system. The python libraries are commonly used to get improved performance of the system. After inputting the text, length is calculated using 'len' function.

```
# Length of text
len(mytext)
output: 607
```

### word\_list=mytext.split()

```
Output: ["'केद्रीय'", 'माध्यमिक', 'शिक्षण', 'मंडळ'तर्फे',
'(सीबीएसई)', 'इयत्ता', 'दहावीचा', 'निकाल', 'आज,', '१५', 'जुलैला',
'जाहीर', 'केला', 'जाणार', 'आहे.', 'निकालाची', 'नेमकी', 'वेळ', 'मात्र',
'बोर्डनि', 'जाहीर', 'केलेली', 'नाही.', 'बारावीचा', 'निकाल', 'सोमवारी,',
'१३', 'जुलैला', 'जाहीर', 'झाला', 'होता.', 'दहावीचा', 'निकाल',
"'सीबीएसई'चे", 'अधिकृत', 'संकेतस्थळ', 'cbse.nic.in', 'येथे',
'दिसेल.', 'यासोबतच', 'निकालासाठी', 'स्वतंत्र', 'पेज', 'असलेल्या',
'cbseresults.nic.in', 'या', 'लिंकवरही', 'तो', 'पाहता', 'येईल.', 'यंदा',
"'सीबीएसई'ने", 'आयव्हीआरएस', 'सुविधा', 'उपलब्ध', 'करून',
'दिली', 'आहे.', 'यासाठी', 'विद्यार्थ्यांना', '०११-२४३००६९९,', '०११-
२८१२७०३०', 'या', 'क्रमांकांवर', 'संपर्क', 'साधावा', 'लागेल.', 'कॉल',
'सुरू', 'असतानाच', 'विचारले', 'गेल्यावर', 'मोबाइलवर', 'आपला',
'रोल', 'नंबर', 'आणि', 'जन्मतारीख', 'टाकल्यानंतर', 'निकाल',
'समजेल.']
```

The next step is tokenization, where the sentences are broken into tokens. The process of tokenization includes splitting the text, where Text.Split() can be used and then the list of all the words is forwarded for next step.

The further step in pre-processing is to remove special characters or symbols in the tokenized document. These characters are searched in the document, and for this we

used Text.Replace()function, which searches for the special characters first and replaces them with white spaces.

```
for char in ' " ' " " ' ~ , / ? ' ' [ ] { } : ; \ | ~ ! @
# $ % ^ & * ( ) _ - = + < > \n ':
Text= mytext.replace(char , ' ')
" 'केद्रीयमाध्यमिकशिक्षणमंडळ'तर्फे (सीबीएसई)
इयत्तादहावीचानिकालआज,१५जुलैलाजाहीरकेलाजाणारआहे.निकाला
चीनेमकीवेळमात्रबोर्डनिजाहीरकेलेलीनाही. बारावीचानिकालसोमवारी,
१३जुलैलाजाहीरझालाहोता.दहावीचानिकाल
'सीबीएसई'चेअधिकृतसंकेतस्थळ cbse.nic.in
येथेदिसेल.यासोबतचनिकालासाठीस्वतंत्रपेजअसलेल्या
cbseresults.nic.in यालिंकवरहीतोपाहतायेईल. यंदा
'सीबीएसई'नेआयव्हीआरएससुविधाउपलब्धकरूनदिलीआहे.
\nयासाठीविद्यार्थ्यांना०११-२४३००६९९, ०११-
२८१२७०३०याक्रमांकांवरसंपर्कसाधावालागेल.कॉलसुरूअसतानाचवि
चारलेगेल्यावरमोबाइलवरआपलारोलनंबरआणिजन्मतारीखटाकल्यानंतर
निकालसमजेल. "
```

We Have to count frequency of each word because the irrelevant words i.e. An empty array is created for storing the count; to calculate this frequency count get () function is used and counter will help to get exact count of each word then.

### for word in word\_list:

d[word]= d.get(word,0)+1

```
output:{"केद्रीय": 1, 'माध्यमिक': 1, 'शिक्षण': 1, 'मंडळ'तर्फे': 1,
'(सीबीएसई)': 1, 'इयत्ता': 1, 'दहावीचा': 2, 'निकाल': 4, 'आज,': 1,
'१५': 1, 'जुलैला': 2, 'जाहीर': 3, 'केला': 1, 'जाणार': 1, ...
```

The Key Value pairs are formed then for feature vector. It gives a list of words and its frequency count in front of that word as shown in the following figure, this step gives feature vector for the input document.

### for key, value in d.items():

word\_freq.append((value,key))

#### Output:

```
[["केद्रीय", 1],
[1, 'माध्यमिक'],
[1, 'शिक्षण'],
[1, "मंडळ'तर्फे"],
[1, '(सीबीएसई)', 1],
[1, 'इयत्ता'],
[2, 'दहावीचा'],
[4, 'निकाल'],
[1, 'आज,'],
[1, '१५'],
[2, 'जुलैला'],
[3, 'जाहीर'],
[1, 'केला'],
[1, 'जाणार'],
[2, 'आहे.'],
[1, 'निकालाची'],
[1, 'नेमकी'],
[1, 'वेळ'],
[1, 'मात्र'].....
```

## IV. CONCLUSION

There is a necessity that the regional language e-content must be focussed for text summarization. This paper gives a spotlight on the regional language of Maharashtra i.e. Marathi. The tools used for processing the Marathi text are in a way effectual, because the efficacy changes depending on the language and tools used for text summarization. The paper highlights the flow of pre-processing by which the Marathi text goes for summarization. In first step, the input file is extracted, then the length of text is calculated, tokenization is performed, end of the sentence is calculated, special symbols are removed, then the frequency count of the word is taken as a statistical value and key value pairs are formed for further processing. We are trying to develop a system which is comparatively more capable and efficient for summarizing Marathi e-News.

## REFERENCES

1. Mr. Shubham Bhosale, Ms. Diksha Joshi, Ms. VrushiBhise, Prof. Rushali A. Deshmukh, "Marathi e-Newspaper Text Summarization Using Automatic Keyword Extraction Technique", International Journal of Advance Engineering and Research Development Volume 5, Issue 03, March -2018.
2. Pooja Bolaj, SharvariGovilkar, "Text Classification for Marathi Documents using Supervised Learning Methods", International Journal of Computer Applications (0975 – 8887), Volume 155 – No 8, December 2016.
3. Virat V. Giri, Dr.M.M. Math and Dr.U.P. Kulkarni, "A Survey of Automatic Text Summarization System for Different Regional Language in India", Bonfring International Journal of Software Engineering and Soft Computing, Vol. 6, Special Issue, October 2016.
4. Prof. Satish Kamble, ShivilMandage, ShubhangiTopale, DipaliVagare, PreranaBabbar, "Survey on Summarization Techniques and Existing Work", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 1 (2017).
5. Anishka Chaudhari1, Akash Dole2, Deepali Kadam3, "Marathi text summarization using neural networks", International Journal of Advance Research and Development, Volume 4, Issue 11, 2019.
6. Deepali K. Gaikwad, Deepali Sawane and C. Namrata Mahender, "Rule Based Question Generation for Marathi Text Summarization using Rule Based Stemmer", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p-ISSN: 2278-8727, PP 51-54, 2018.
7. Yogeshwari V. Rathod, "Extractive Text Summarization of Marathi News Articles", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 07, July 2018.
8. Shraddha A. Narhari, RajashreeShedge, "Text Categorization of Marathi Documents using Modified LINGO", IEEE, 2017
9. Jaydeep Jalindar Patil, Prof. NagarajuBogiri, "Automatic Text Categorization-Marathi documents", International Conference on Energy Systems and Applications (ICESA 2015), IEEE, 2015.
10. Prakhar Sethi, Sameer Sonawane, SaumitraKhanwalker, R. B. Keskar, "Automatic Text Summarization of News Articles", International Conference on Big Data, IoT and Data Science (BIG) Vishwakarma Institute of Technology, Pune, Dec 20-22, IEEE, 2017
11. N. Dangre, A. Bodke, A. Date, S. Rungta, S.S. Pathak, "System for Marathi news clustering", 2nd International conference on Intelligent computing, communication & convergence, bhubaneswar, ELSEVIER, 2016.
12. Apurva D. Dhawale, Sonali B. Kulkarni, Vaishali Kumbhakarna, "Survey of Progressive Era of Text Summarization for Indian and Foreign Languages Using Natural Language Processing", ICIDCA 2019, LNDECT 46, pp. 654–662, Springer Nature Switzerland, AG, 2020.
13. E. Lloret and M. Palomar, "Text summarization in progress: a literature review," in Springer, no. April 2011, pp. 1–41, Springer, 2012.
14. Tarun B. Mirani and SreelaSasi, "Two-level Text Summarization from Online News Sources with Sentiment Analysis", International Conference on Networks & Advances in Computational Technologies (NetACT) ,20-22 July 2017, Trivandrum, IEEE, 2017.
15. Vaishali Kalra, Dr. Rashmi Aggarwal, "Importance of Text Data Preprocessing & Implementation in RapidMiner", Proceedings of the First International Conference on Information Technology and

- Knowledge Management pp. 71–75. ICITKM, ISSN 2300-5963 ACSIS, Vol. 14, New Delhi, 2017.
16. Sheetal Shimpikar, Sharvari Govilkar, "Abstractive Text Summarization using Rich Semantic Graph for Marathi Sentence", JASC: Journal of Applied Science and Computations Volume V, Issue XII, ISSN NO: 1076-5131, December/2018.
17. Jovi D'silva, Dr. Uzzal Sharma, "Automatic Text Summarization Of Indian Languages: A Multilingual Problem", Journal of Theoretical and Applied Information Technology Vol.97. No 11, 15th June 2019.
18. Poonam Kolhe, Prof. Ashish Kumbhare, "Optimizing Accuracy of Document Summarization Using Rule Mining", International Journal of Computer Science and Mobile Computing, Vol.6 Issue.6, pg. 207-216, June- 2017.
19. Umakant Dakulge, S. C. Dharmadhikari, "Automated Text Summarization: A Case Study for Marathi Language", Data Mining and Knowledge Engineering, CIIT, Vol 6, No 3 (2014).
20. Mamatha Balipa, Dr. Balasubramani R, Harolin Vaz, Christina Shilpa Jathanna, "Text Summarization For Psoriasis Of Text Extracted From Online Health Forums Using Textrank Algorithm", International Journal Of Engineering & Technology, 7 (3.34) (2018) 872-873, 18 September 2018.
21. Chirantana Mallick, Ajit Kumar Das, Madhurima Dutta, Asit Kumar Das And Apurba Sarkar, "Graph-Based Text Summarization Using Modified Textrank", J. Nayak Et Al. (Eds.), Soft Computing In Data Analytics, Advances In Intelligent Systems And Computing 758, Springer Nature Singapore Pte Ltd. 2019.
22. 10| Reda Elbarougy, Gamal Behery, Akram El Khatib, "Extractive Arabic Text Summarization Using Modified Pagerank Algorithm", Egyptian Informatics Journal 21, 73–81, Science Direct, Elsevier, (2020).
23. Ahmed Elrefaiy, Ahmed Rafat Abas, Ibrahim Elhenawy, "Review Of Recent Techniques For Extractive Text Summarization", Journal Of Theoretical And Applied Information Technology 15th December 2018. Vol.96. No 23, Issn: 1992-8645, Jatit & Lls, 2005.
24. Rasim Alguliev, Ramiz Aliguliyev, "Evolutionary Algorithm for Extractive Text Summarization", Intelligent Information Management, 1, 128-138, Scientific Research, SciRes, 2009.
25. Kalliath Abdul Rasheed Issam, Shivam Patel, Subalalitha C. N., "Topic Modeling Based Extractive Text Summarization", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-6, April 2020.
26. Siddhant Upasani, Noorul Amin, Sahil Damania, Ayush Jadhav, A. M. Jagtap, "Automatic Summary Generation using TextRank based Extractive Text Summarization Technique", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, Volume: 07 Issue: 05 May 2020.
27. Yash Asawa, Vignesh Balaji, Ishan Isaac Dey, "Modern Multi-Document Text Summarization Techniques", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-9 Issue-1, May 2020.

## AUTHORS PROFILE



**Ms. Apurva D. Dhawale** completed M.phil in Computer Science in 2015 from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India. Currently she is pursuing her Ph.D. in Computer Science from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India. She has 9 years of teaching experience in Dr. G. Y. Pathrikar College of CS & IT, MGM University,

Aurangabad and published 9 papers reputed international journals including Scopus, Elsevier, Springer. Her research interest areas are Natural Language Processing & Biometric Image Processing.



**Dr. Sonali B Kulkarni** Completed her Master of Science from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India with First in the order of merit in year 2002. She has also completed Ph.D in Computer Science from Dr. BAM University, Aurangabad and currently working as Assistant Professor in Department of Computer Science and IT,

Dr.BAMUniversity, Aurangabad since 2008. She has 18 years of teaching and Research Experience. She is a life member of IETE, ISCA, Grant and IJENG She has published more than 30 research papers in reputed international journals including Scopus, Thomson Reuters, ICI, Web of Science and conferences including IEEE, Elsevier, Springer. Her main research work focuses on Remote Sensing & GIS, Brain Computer Interface, Natural Language Processing and Linguistics.



**Ms.Vaishali M. Kumbhakarna** completed Master of Science from Dr.Babasaheb Ambedkar Marathwada University, Aurangabad, India with First class in year 2013. And M.C.M. from Pune University in 2001.She has also completed M.Phil in Computer Science from Dr.BAMUniveristy, Aurangabad in the year 2016 and Currently she is pursuing her Ph.D. in Computer Science from Dr.Babasaheb Ambedkar Marathwada University,

Aurangabad, India. currently working as Assistant Professor in Department of Computer Science and IT, M.G.M University, Aurangabad. since 2008. She has 12 years of teaching experience. She has published 8 research papers in reputed international journals including Scopus, Elsevier, Springer. Her main research work focuses on, biomedical digital image processing and Natural Language Processing