How to Cite:
Prakash, N. C. P., Narasimhaiah, A. P., Nagaraj, J. B., Pareek, P. K., Maruthikumar, N. B., & Manjunath, R. I. (2022). Implementation of NLP based automatic text summarization using spacy. *International Journal of Health Sciences*, *6*(S5), 7508–7521.
https://doi.org/10.53730/ijhs.v6nS5.10574

Implementation of NLP based automatic text summarization using spacy

Nayana Cholanayakanahalli Prakash
Post Graduation Student, Master of Technology, Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru, Karnataka, India; Pincode: 560091
Corresponding author email: nayanaewit@gmail.com

Achyutha Prasad Narasimhaiah
Professor, Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru, Karnataka, India; Pincode: 560091
Email: achyuth001@gmail.com

Jagadeesh Bettakote Nagaraj
Assistant Professor, Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru, Karnataka, India; Pincode: 560091
Email: jagadeeshbn001@gmail.com

Piyush Kumar Pareek
Professor, Department of Computer Science and Engineering & Head of IPR Cell, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India; Pincode: 560064
Email: piyush.kumar@nmit.ac.in

Nalini Bpalya Maruthikumar
Assistant Professor, Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru, Karnataka, India; Pincode: 560091
Email: nalinibm03@gmail.com

Ramya Iyaravally Manjunath
Assistant Professor, Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru, Karnataka, India; Pincode: 560091
Email: ramyaim31@gmail.com

***Abstract*---**The amount of data on the Internet has increased exponentially over the past decade. Therefore, we need a solution that converts this massive amount of raw information into useful information that the human brain can understand. One such common technique in research that helps when dealing with large amounts of data is text summarization. Automatic summarization is a well-known

International Journal of Health Sciences ISSN 2550-6978 E-ISSN 2550-696X © 2022.
Manuscript submitted: 9 April 2022, Manuscript revised: 18 June 2022, Accepted for publication: 27 July 2022

approach to reduce documents to key ideas. This works by storing important information by creating a shortened version of the text. Text summaries are divided into extraction and abstraction methods. The extraction summary method minimizes the summarization burden by selecting a subset of relevant sentences from the actual text. There are many methods, but researchers specializing in natural language processing (NLP) are particularly attracted to the extraction method. The meaning of the sentence is calculated using linguistic and statistical features. In this work, extractive and abstract methods for summarizing texts were examined. This white paper uses a spacey algorithm to analyze the above methods, resulting in fewer iterations and a more focused summary.

***Keywords*---**Empirical Methods, Text Summarization, Extraction, Abstraction, Reinforcement Learning, Supervised, Unsupervised, NLP, Spacy Algorithms.

Introduction

Summarizing large texts remains an open topic in natural language processing. Automatic text summarization is used to summarize large documents. Text summary is the process of using software to simplify a text document and create an overview of the gist of the original document. Text summarization is a technique that software uses to shrink a text document to create a summary or synopsis of the original document. Summary is performed to highlight important parts of the text. Text summaries can be categorized based on the input type. Single low-input document in a text context. In such cases, a simple summary model is created. Several documents, the input of which can be relatively long. The more text you have, the more semantic links are generated, which adds to the complexity here. Summarizers can be categorized as general based on their purpose. It is abstract and the model forms its own phrases and sentences to provide a more consistent summary. In this case, the model processes the input without bias or prior knowledge. Domain specific. The model uses domain information to create a more accurate summary based on known facts. Query based. The summary contains only known answers to natural language questions about the input text. Based on output type, totalizers can be categorized as follows: Extract. Select a key phrase from the input text to create a summary that humans generate. Creating an abstract summary is generally a more complex task than the extraction method. Therefore, apart from recent advances in the use of neural networks, driven by advances in neural machine translation and sequence models, they are still far from reaching the human level. Text summarization applications include media monitoring, search engine marketing, internal document workflows, financial research, social media marketing, and assisting people with disabilities.

Method

This segment reviews a number of excellent text summaries as shown in Figure 1.

Unsupervised Extracts

Unsupervised summarization techniques mean creating summaries from a given document without using previously identified groups or classifications. There are three ways to do this: first graph-based, latent variables, and finally term frequency. These are easy to implement and give satisfactory results. Some of the surveys conducted are listed below. Hernándecetal. presented a solution for selecting sentences in an extract summary using KMeans clustering. This is a big disadvantage. The first step is to remove stop words, hyphens and extra spaces. This is called input text preprocessing.

The next step is to use n-grams to select features and use boolean weight (BOOL), term frequency (TF), inverse document frequency (IDF), or TFIDF to find weights. The next step is to apply KMeans to the clustering set. KMeans is an iterative process that plots values up to the nearest centroid (the average of all values) and then calculates a new centroid. In the proposed method, the first sentence is considered as the baseline and the similarity between the sentences is plotted using Euclidean distance. After clustering is done using the K-cluster, the set that is closest to the centroid (aka the most representative set) is selected. The proposed method achieves better results than other prior art methods.

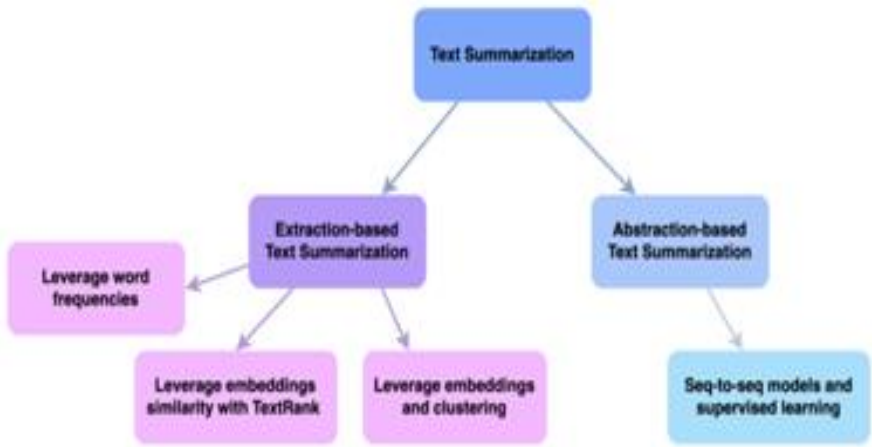

Figure 1: Classification of text summaries

Joshi et al. proposed an unsupervised framework for extracting text summarization of a single document called SummCoder. After preprocessing, SummCoder uses a jump thinking model to transform the sentence into a fixed-length vector. To create a summary, sentences are selected considering three scores: Sentence Content Relevance Metric (scoreContR), Sentence Novelty Metric (scoreNov), and Sentence Position Relevance Metric (scorePosR). After all points have been calculated, the final score and relative score will be calculated. Finally, summaries can be generated by first sorting by descending relative rank and then by appearance in the input text.

El-Kassas et al. announced a single document and graph-based extraction system called EdgeSumm. In the proposed method, a pretreatment and a lemming are carried out first. Then a graph of the textual representation is created with nouns as nodes and non-noun words as edges. There are "S#" and "E#" nodes that indicate the beginning and end of the sentence. For each node, the weight is calculated by counting the number of times it occurs. When choosing a sentence, it is assumed that all nouns represent different topics. First, search for the most common words and phrases and make a list of selected nodes and edges. To select the source and destination nodes, the score must be higher than the average score for all nodes. To select an edge, you must select both the source node and the target node. If the candidate summary (the summary generated by the algorithm) exceeds the user limit, the candidate summary will be scored and ranked in ascending order. Then apply KMeans clustering to group similar statements and select the top statements from each cluster to create the final summary.

Zheng & Lapata, proposed a position-enhanced centrality-based summary (PacSumm). Uses a graph-based rank algorithm. A set is a node and an edge shows the relationship between the nodes. A bi-directional encoder representation of Transformers (BERT) was used to map the set. The BERT pre-training has two tasks. The first task is masked language modeling, where the sentences are tokenized with the left and right sentences. The second task is sentence identification, which identifies the relationship between two sentences. To fine-tune the BERT, five negative samples are reported for each positive sample. After finding the representation of all sets, use the pairwise dot product to create an unnormalized matrix. Sets are selected based on this matrix.

Vanetiketal proposed a weighted compression model to extract important information from text. In the proposed model, this is done by reducing the sentence by repeatedly deleting the elementary discourse unit (EDU). First, each word is given a non-negative weight. Weights are assigned using the Gillic and Favre and McDonald extraction models. The next step is to select and delete the EDU. The list of EDUs is created using the constituency-based syntax tree. Excluded from the list are EDUs, which may make an ungrammatical statement when deleted. Everything else is removed and the "important" EDU weights are calculated and sorted. During summary creation, EDUs are selected based on maximum weight-to-cost ratio and cannot exceed the length of the summary.

Ozsoy & Alpaslan introduced Latent Semantics (LSA) for text summarization. This is an algebraic static method for finding hidden logical patterns between words and sentences. The input matrix is created to display the text. Rows represent words and columns represent sentences. The cell shows the TF-IDF value of the word. Single Value Decomposition (SVD) is used to model the relationship between words and sentences. The result of the SVD is useful for selecting instructions using the cross method. The set with the longest vector is selected.

Discussion

A. NLP and knowledge base

The study of natural language processing (NLP) dates from the 1950s. NLP understand and processes languages using sentence grammar, ontology, language models, analysis trees and similar methods. NLG (Natural Language Generation) does the opposite and generates natural language from machine rendering. NLP / NLG-based summaries are sometimes referred to as "semantics-based or ontology-based" (Allahyari et al., 2017) rather than "knowledge-based. Before the advent of deep learning systems, NLP and ontology-based solutions have been the most common ways to do abstract transformations. For example, sentences can be grouped by mechanical conjunction rules. This type of abstraction is primarily grammatical and may not provide integration of document ideas. Think of this "abstract light".

Some researchers have combined NLP with deep learning to "encode" "linguistic information" such as Part-of-Speech (POS) tags and Named Entity Extraction (NER) tags as a lexical function as part of an encoder-decoder neural network. . (Zhou, Yang, Wei, Tan & Bao, 2017). I agree with Allahyari and others. (2017) "The step to building a more accurate summarization system is to combine the summarization method with knowledge-based and semantics-based ontology-based summarization." The trend that can be seen in the comparison matrix is away from NLP and towards deep learning

B. Spacy

SpaCy is a free open-source natural language processing library in Python programming languages. Used primarily in production software development, spaCy also supports deep learning workflows via PyTorch and TensorFlow statistical models. SpaCy provides fast and accurate parsing, named entity recognition, and easy access to word vectors. You can use the default word vector or replace it with another word vector. SpaCy also features tokenization, sentence boundary detection, part-of-speech tagging, parsing, built-in word vectors, and very accurate placement on the original string.

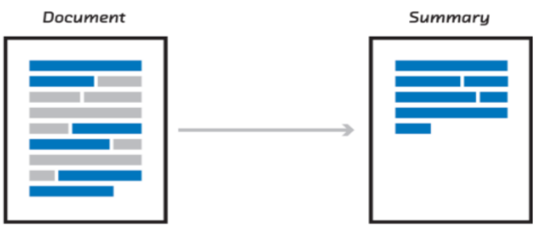

Fig 2: Text Summarization

Text summarization procedure:

1. Text cleaning: Remove stop words, punctuation, and lowercase letters in words.
2. Working tokenization: Tokenize all words from a sentence.

3. Table of Word Frequency each word Frequency is calculated and divide maximum frequency with each frequency to get the normalized word frequency count.
4. Sentence tokenization: As per Frequency of the sentence summary will done.

```
nlp = spacy.load('en_core_web_sm')
doc= nlp(text)
tokens=[token.text for token in doc]
print(tokens)
```

```
sentence_scores = {}
for sent in sentence_tokens:
    for word in sent:
        if word.text.lower() in word_frequencies.keys():
            if sent not in sentence_scores.keys():
             sentence_scores[sent]=word_frequencies[word.text.lower()]
            else:
             sentence_scores[sent]+=word_frequencies[word.text.lower()]
```

The other spacy application we are investigating consists of automatic summarization sentence extraction. In a way, you can think of the sentence extraction problem as similar to keyword extraction. Both applications aim to identify sequences that are more "representative" of a given text. In keyword extraction, candidate text units consist of words, while sentence extraction processes the entire sentence. Text Rank proves suitable for this type of application as it accepts a recursively computed per-text ranking based on information gleaned from the entire text.

The evaluation according to the essential approach is often carried out by comparing the results of the machine overview with the expert's ideal overview. Substantial assessments are made by measuring grammar, verbosity, and consistency. This measurement is very rare because it is a very rare study that poses grammar and consistency issues. Find out how many ideal sentences there are in the automated machine summary and evaluate them from a sentence extraction perspective. The evaluation was carried out using the measurement methods Precision, Recall and f-Score / f-Measure. To evaluate the content, compare the actual words of the sentence, not the entire sentence. The evaluation was made by measuring the similarity of rouge, pyramid and cosine. The benefit of scoring is that human extracts can be compared to an automatic summarization engine

Use a human summary with new sentences and paraphrases. Another assessment is a task-based approach that measures the performance of the automatic summary engine by using the summary for specific tasks such as answering questions and document categories.

Results and Discussions

Architecture of the Project**:**

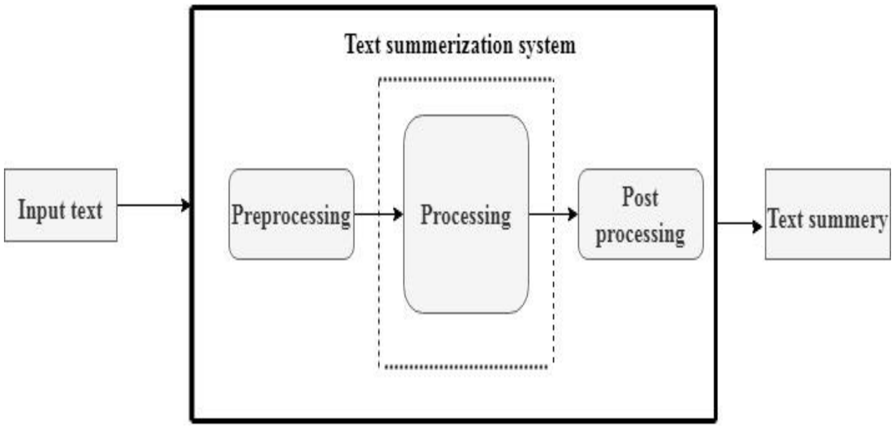

Fig 3: Architecture of the Project

The project architecture is shown in the figure 3 above. As you can see, the text document is uploaded to the application first. The text document is then preprocessed, including the removal of stop words and punctuation, by finding the word Frequency and the Sentence Frequency. Finally, create a text summary.

Document pre-processing

Due to the Excess sources of information in today's world, the input documents we receive may not be in the correct English format, which may contain audio. Sounds include various special characters, unwanted spaces, newlines, stops, and more. Therefore, perform the following tasks on the input file to get only the useful parts of the document.
Step 1: All line breaks are removed.
Step 2: All corner brackets and special numbers are removed.
Step 3: All commas, extra spaces and repeating sentences are removed.

Removal of stop words

In this step it will remove all subtitles from your input according to your native language. These stop words do not provide reliable information about a particular context. It does not convey any information about this emotion as it builds a collection of emotions like "is", "am", "who" to create an illustration.

Tokenization

Previously, sentences were split into several words. Basically, this token model is used to do the activity in the form of a pipelined NLP natural language processing process. This is useful at two stages, word level and sentence level. The first is a standard word mark that restores a set of words in a given sentence.

Extraction of important sentences

We need a way to verify the value of the text in the scroll. The following calculations are performed to extract the key phrase from the document and the same is shown in figure 4 & 5.
Step 1: Frequency of all words in the previously reviewed text is calculated.
Step 2: The weight of each word by dividing the frequency of the words by the maximum frequency is calculated.
Step 3: Review all key phrases for the specified input.
Step 4: The sentence's score by adding the weighted frequencies of the words contained in the sentence is calculated.
Step 5: Sort the set's token list in descending order based on points
Step 6: Remove the "n" statement from the token list

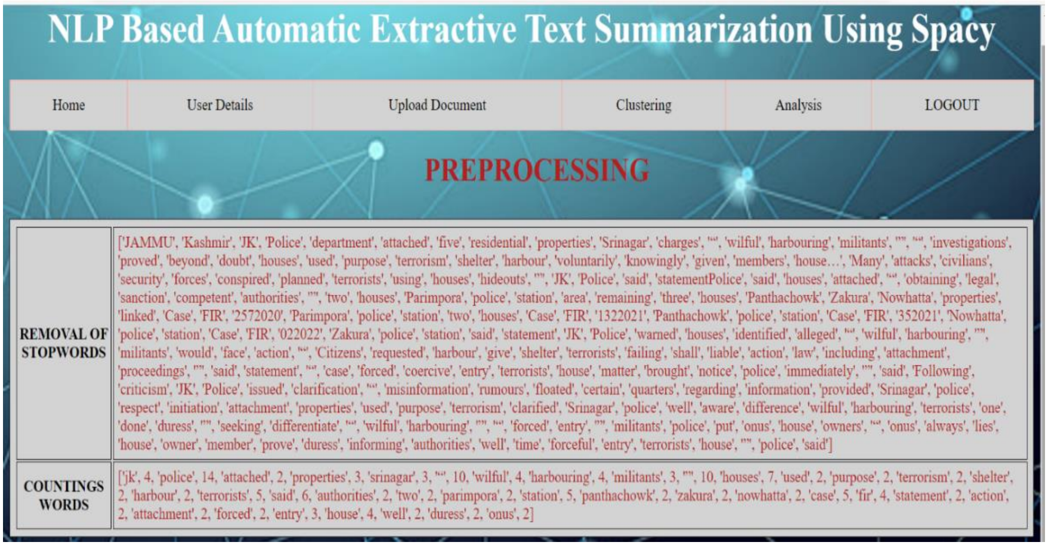

Figure 4: Preprocessing

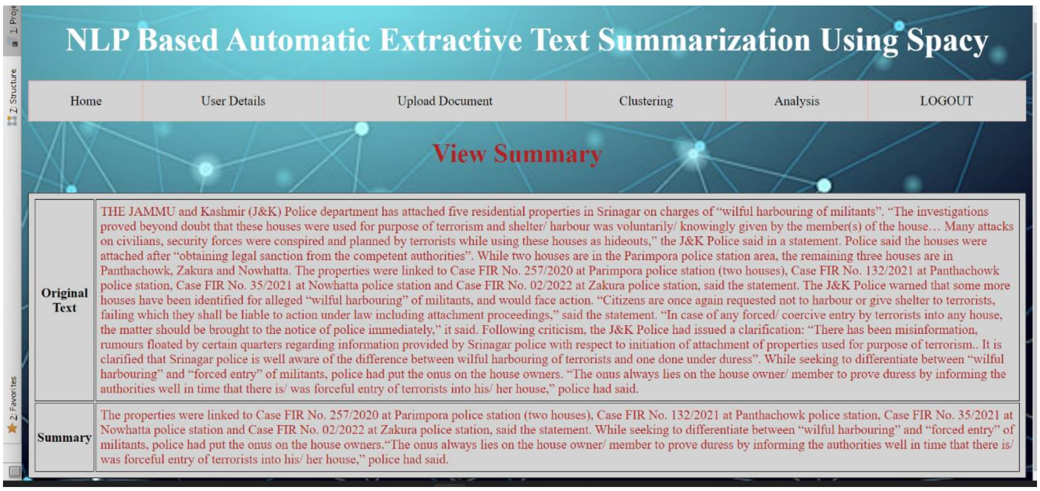

Fig 5: Viewing the Summary

Conclusion

Automatic text summarization is an interesting academic topic with a wide range of commercial applications. By reducing huge amounts of information into short bursts, summaries are useful in a variety of downstream applications such as news summaries, reporting, news summaries and headlining. There are two types of aggregation algorithms that are most commonly used. The extraction summary method starts with rearranging and copying the passages from the source material. Second, the abstract summary approach creates new phrases by rephrasing or inserting terms not found in the original text. Most of the research to date has been extractive due to the difficulty of abstract summarization. The extraction approach is more convenient as it guarantees grammar and accuracy by copying a large chunk of text from the source document. On the other hand, advanced skills such as paraphrasing, generalizing, and assimilation of real-world knowledge are only possible with abstract frameworks and are required for high-quality summarization. Despite the fact that abstract summarization is a more difficult task, there has been some success thanks to recent advances in deep learning.

Acknowledgments

I extend my deep sense of sincere gratitude to **Dr. Channakesavalu K**, Principal, East West Institute of Technology, Bengaluru, for having permitted to carry out the survey on **“NLP Based Automatic Text Summarization using Spacy”** successfully.
I express my heartfelt sincere gratitude to my guide **Dr. Achyutha Prasad N**, Head, Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru for his valuable guidance, encouragement and suggestions.

I would like to express my sincere thanks to my internal co-guide **Jagadeesh B N**, Assistant Professor, Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru for his valuable guidance, encouragement and suggestions.

I would like to thank all the Teaching, Technical faculty and supporting staff members of Department of Computer Science and Engineering, East West Institute of Technology, Bengaluru, for their valuable suggestions and support.

Finally, I would like to thank my **Parents** for their support.

References

A. P. N and C. D. Guruprakash, "A Relay Node Scheme for Energy Redeemable and Network Lifespan Enhancement," 2018 4th International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Mangalore, India, 2018, pp. 266-274.

Abbasi-ghalehtaki, R., Khotanlou, H., and Esmaeilpour, M. (2016). Fuzzy evolutionary cellular learning automata model for text summarization. Swarm and Evolutionary Computation, 30:11–26.

Abdi, A., Shamsuddin, S. M., and Aliguliyev, R. M. (2018). Qmos: Query-based multi-documents opinion-oriented summarization. Information Processing & Management, 54(2):318–338. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

Achyutha Prasad, N., Guruprakash, C.D., 2019. A relay mote wheeze for energy saving and network longevity enhancement in WSN. International Journal of Recent Technology and Engineering 8, 8220–8227. doi:10.35940/ijrte.C6707.

Achyutha Prasad, N., Guruprakash, C.D., 2019. A relay node scheme of energy redeemable and network lifespan enhancement for wireless sensor networks and its analysis with standard channel models. International Journal of Innovative Technology and Exploring Engineering 8, 605–612.

Achyutha Prasad, N., Guruprakash, C.D., 2019. A two hop relay battery aware mote scheme for energy redeemable and network lifespan improvement in WSN. International Journal of Engineering and Advanced Technology 9, 4785–4791. doi:10.35940/ijeat.A2204.109119.

Achyutha, P. N., Hebbale, S., & Vani, V. (2022). Real time COVID-19 facemask detection using deep learning. International Journal of Health Sciences, 6(S4), 1446–1462. https://doi.org/10.53730/ijhs.v6nS4.6231.

Automatic text summarization using reinforcement learning with embedding features. In Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers), volume 2, pages 193–197.

Chetana Srinivas, Ambrish G, Bharathi Ganesh, Anitha Ganesh, Dhanraj, Kiran M, “Logistic Regression Technique for Prediction of Cardiovascular Disease”, International Conference on Intelligent Engineering Approach,(ICIEA) India, 12th February 2022.

Chetana Srinivas, Ambrish G, Supritha N, Bharathi G, Anitha G, “Survey on Recent Trends in Machine Learning and Deep Learning in Healthcare”, International Conference on Recent Trends in Machine Learning and Computing System,(RTMCS) India, 17th -18th December 2021.

Chetana Srinivas, Nandini Prasad K S,"A Comparative study on Medical Image Processing Using Big Data Analytics Frameworks”, 2018 Third International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEEECCOT), Mysuru, India, 2018.

Chetana Srinivas, Nandini Prasad K S,” A Comparative Study on Different Types of Image Pre-processing Methods for Noise Removal”, Internal Journal of Computing, Communication & Networking (IJCCN), ISBN: 2319-2720, Vol.7, Issue 2, April 2018.

Chetana Srinivas, Nandini Prasad K. S., Mohammed Zakariah, Yousef Ajmi Alothaibi , Kamran Shaukat , B. Partibane, and Halifa Awal, “Deep Transfer Learning Approaches in Performance Analysis of Brain Tumor Classification Using MRI Images”, Hindawi Journal of Healthcare Engineering Volume 2022, Article ID 3264367, 17 pages https://doi.org/10.1155/2022/3264367.

Daum´e III, H. and Marcu, D. (2004). A tree-position kernel for document compression. In Proceedings of DUC2004.

Dr.Balakrishna R, Piyush Kumar Pareek et al, ’Data Mining for Healthy Tomorrow with the implementation of Software Project Management technique’, Springer AISC Series/ SCOUPS INDEXED JOURNAL, Paper Id : IT -187-ICPCIT2015, June 2015.

Dr.Balakrishna R, Piyush Kumar Pareek et al, 'Study on Six Sigma approach to improve the quality of process outputs in business processes in Small & Medium Level Software Firms' Springer AISC Series/ SCOUPS INDEXED JOURNAL, Paper Id : IT -221-ICPCIT2015.

Dr.Piyush Kumar Pareek et al, 'A survey on approaches for predicting performance of students',International Journal of Engineering Research and Science, ISSN No.2395-6992 Paper Id:IJOER-Jun-2016-25.

Dr.Piyush Kumar Pareek et al, 'A survey on Long term product planning and requirements prioritization to customer value creation', International Journal of Engineering Research and Science, ISSN No.2395-6992 Paper Id: IJOER-Jun-2016-27.

Dr.Piyush Kumar Pareek et al, 'Education Data Mining –Perspectives of Engineering Students ', International Journal of Innovative Research in Computer Science & Technology (IJIRCST), ISSN: 2347-5552, Volume-4, Issue-5, September-2016.

Edmundson, H. P. (1969). New methods in automatic extracting. Journal of the ACM, 16(2):264–285.

Hebbale, S., Marndi, A., Achyutha, P. N., Manjula, G., Mohan, B. R., & Jagadeesh, B. N. (2022). Automated medical image classification using deep learning. International Journal of Health Sciences, 6(S5), 1650–1667. https://doi.org/10.53730/ijhs.v6nS5.9153.

Hebbale, S., Marndi, A., Manjunatha Kumar, B. H., Mohan, B. R. ., Achyutha, P. N., & Pareek, P. K. (2022). A survey on automated medical image classification using deep learning. International Journal of Health Sciences, 6(S1), 7850–7865. https://doi.org/10.53730/ijhs.v6nS1.6791.

Hovy, E. and Lin, C. Y. (1999). Automated text summarization in summarist. In Mani, I. and Maybury, M. T., editors, Advances in Automatic Text Summarization, pages 81–94. MIT Press.

Jipeng, T., Neelagar, M. B., & Rekha, V. S. (2021). Design of an embedded control scheme for control of remote appliances. Journal of Advanced Research in Instrumentation and Control Engineering, 7(3 & 4), 5-8.

Kadakadiyavar, S., Prasad, A. N., Pareek, P. K., Vani, V., Rekha, V. S., & Nirmala, G. (2022). Recognition efficiency enhancement of control chart pattern using ensemble MLP neural network. International Journal of Health Sciences, 6(S3), 4295–4306. https://doi.org/10.53730/ijhs.v6nS3.6851.

Kalshetty, J. N., Achyutha Prasad, N., Mirani, D., Kumar, H., & Dhingra, H. (2022). Heart health prediction using web application. International Journal of Health Sciences, 6(S2), 5571–5578. https://doi.org/10.53730/ijhs.v6nS2.6479.

Knight, K. and Marcu, D. (2000). Statistics-based summarization - step one: Sentence compression. In AAAI/IAAI, pages 703–710.

Kogilavani, A. and Balasubramanie, P. (2010). Clustering based optimal summary generation using genetic algorithm. In Communication and Computational Intelligence (INCOCCI), 2010 International Conference on, pages 324–329. IEEE. Lee, G. H. and Lee, K. J. (2017).

Kupiec, J., Pedersen, J., and Chen, F. (1995). A trainable document summarizer. In Proceedings SIGIR '95, pages 68–73, New York, NY, USA.

Lebanon, G. (2006). Sequential document representations and simplicial curves. In Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence.

Luhn, H. P. (1958). The automatic creation of literature abstracts. IBM Journal of Research Development, 2(2):159–165.

Mani, I. and Bloedorn, E. (1997). Multi-document summarization by graph search and matching. In AAAI/IAAI, pages 622–628.

Manjunatha Kumar, B. H., Achyutha , P. N., Kalashetty, J. N., Rekha, V. S., & Nirmala, G. (2022). Business analysis and modelling of flight delays using artificial intelligence. International Journal of Health Sciences, 6(S1), 7897–7908. https://doi.org/10.53730/ijhs.v6nS1.6735.

Mehdi Allahyari and Krys Kochut. 2015. Automatic topic labeling using ontology-based topic models. In Machine Learning and Applications (ICMLA), 2015 IEEE 14th International Conference on. IEEE, 259–264.

Mr. Piyush Kumar Pareek, Dr. A. N. Nandakumar, Lean software development Survey on Agile and Lean usage in small and medium level firms in Bangalore, International Journal of Advanced Research in Computer Science and Software Engineering , Volume 4, Issue 12, December 2014 , ISSN: 2277 128X .pp 1-7 Impact Factor : 2.08.

Mr.Piyush Kumar Pareek, Dr. A. N. Nandakumar, 'Lean software development Survey on Benefits and challenges in Agile and Lean usage in small and medium level firms in Bangalore' , International Journal of Advanced Research in Computer Science and Software Engineering , Volume 4, Issue 12, December 2014 , ISSN: 2277 128X .pp 1-11.

N. A. Prasad and C. D. Guruprakash, "An ephemeral investigation on energy proficiency mechanisms in WSN," 2017 3rd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Tumkur, 2017, pp. 180-185.

N. G and G. C. D, "Unsupervised Machine Learning Based Group Head Selection and Data Collection Technique," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), 2022, pp. 1183-1190, doi: 10.1109/ICCMC53470.2022.9753995.

Narayan, S., Cohen, S. B., and Lapata, M. (2018). Ranking sentences for extractive summarization with reinforcement learning. arXiv preprint arXiv:1802.08636.

Oufaida, H., Nouali, O., and Blache, P. (2014). Minimum redundancy and maximum relevance for single and multidocument arabic text summarization. Journal of King Saud University-Computer and Information Sciences, 26(4):450– 461.

Parveen, D., Mesgar, M., and Strube, M. (2016). Generating coherent summaries of scientific articles using coherence patterns. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 772–783.

Piyush Kumar Pareek & Dr. A. N. Nandakumar, 'To Implement Lean software development frame- work for minimizing waste in terms of non-value added activities', Research Publishing, Jain University ICISTSI-15 , Innovative Partners for Publishing Solutions, Singapore (May 2015).

Piyush Kumar Pareek & Dr.A.N.Nandakumar, 'Failure Mode Effective Analysis of Requirements Phase in small software Firms', Paper ID: ICSTM/YMCA/2015/292, International Conference on Science, Technology and Management (ICSTM-2015). International Journal of Advance Research in Science and Engineering (IJARSE, ISSN- 2319-8354, Impact Factor- 1.142) [www.ijarse.com], Special Issue Jan2015.

Piyush Kumar Pareek & Dr.A.N.Nandakumar, 'Identifying Wastes in software, International Journal of Engineering Studies and Technical Approach'. January Issue 2015.
Piyush Kumar Pareek , Dr.Praveen Gowda , et al 'Ergonomics in a Foundry in Bangalore to improve productivity',International Journal of Engineering and Social Science , ISSN: 2249- 9482 ,Volume 2,Issue 5 (May 2012) , pp 1-6.
Piyush Kumar Pareek , Dr.Praveen Gowda, et al 'FMEA Implementation in a Foundry in Ban- galore to Improve Quality and Reliability', International Journal of Mechanical Engineering and Robotics Research, ISSN :2278-0149,Volume 1,Issue 2(June 2012),pp 81-87.
Piyush Kumar Pareek et al, 'Survey on Challenges in Devops ', International Journal of Innovative Research in Computer Science & Technology (IJIRCST), ISSN: 2347-5552, Volume-4, Issue-5, September-2016.
Piyush Kumar Pareek, Dr. A. N. Nandakumar, et al 'Methodology and Functioning of Project Management Techniques in Agile Software Development Process', International Journal of Research in IT, Management and Engineering, ISSN: 2249-1619, Volume2, Issue12 (December2012), pp 76-85.
Piyush Kumar Pareek, Dr. Vasanth Kumar S A , et al 'Reduction of Cycle Time By Implementation of a Lean Model Carried Out In a Manufacturing Industry', International Journal of Engineering and Social Science , ISSN: 2249-9482,Volume 2, Issue 5, pp 114-123.
Piyush Kumar Pareek, Dr.Vasanth Kumar S A , et al 'Implementation of a Lean Model for Carrying out Value Stream Mapping in a Manufacturing Industry', International Journal of Mechanical Engineering and Robotics Research, ISSN :2278-0149,Volume 1,Issue 2(June 2012),pp 88-95.
Pooja Chopra, Vijay Suresh Gollamandala, Ahmed Najat Ahmed, S. B. G. Tilak Babu, Chamandeep Kaur, N. Achyutha Prasad, Stephen Jeswinde Nuagah, " Automated Registration of Multiangle SAR Images Using Artificial Intelligence & quot, Mobile Information Systems, vol. 2022, Article ID 4545139, 10 pages, 2022. https://doi.org/10.1155/2022/4545139.
Prasad N. Achyutha, Sushovan Chaudhury, Subhas Chandra Bose, Rajnish Kler, Jyoti Surve, Karthikeyan Kaliyaperumal, "User Classification and Stock Market-Based Recommendation Engine Based on Machine Learning and Twitter Analysis", Mathematical Problems in Engineering, vol. 2022, Article ID 4644855, 9 pages, 2022. https://doi.org/10.1155/2022/4644855.
R. V S and Siddaraju, "Defective Motes Uncovering and Retrieval for Optimized Network," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), 2022, pp. 303-313, doi: 10.1109/ICCMC53470.2022.9754109.
Rautray, R. and Balabantaray, R. C. (2017). An evolutionary framework for multi document summarization using cuckoo search approach: Mdscsa. Applied Computing and Informatics. Salton, G. and McGill, M. J. (1986). Introduction to modern information retrieval. Sanchez-Gomez, J. M., Vega-Rodr´ıguez, M. A., and Perez, ´ C. J. (2018). Extractive multi-document text summarization using a multi-objective artificial bee colony optimization approach. Knowledge-Based Systems, 159:1–8.
Rekha VS, Siddaraju., "An Ephemeral Analysis on Network Lifetime Improvement Techniques for Wireless Sensor Networks", International Journal of Innovative Technology and Exploring Engineering, vol. 8, issue 9, 2278-3075, pp. 810–814, 2019.

Sagar, Y.S. and Achyutha Prasad, N., CHARM: A Cost-Efficient Multi-Cloud Data Hosting Scheme With High Availability, International Journal for Technological Research In Engineering, Volume 5, Issue 10, June-2018, ISSN (Online): 2347 – 4718.

Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2022). Post-pandemic health and its sustainability: Educational situation. International Journal of Health Sciences, 6(1), i-v. https://doi.org/10.53730/ijhs.v6n1.5949

Towards a unified approach to simultaneous single-document and multi-document summarizations. In Proceedings of the 23rd international conference on computational linguistics, pages 1137–1145. Association for Computational Linguistics.

Udit Shinghal, Yashwanth A V Mowdhgalya, Vaibhav Tiwari, Achyutha Prasad N "Centaur - A Self-Driving Car" International Journal of Computer Trends and Technology 68.4 (2020):129-131.

Udit Shinghal, Yashwanth A V Mowdhgalya, Vaibhav Tiwari, Achyutha Prasad N "Home Automation using HTTP and MQTT Server" International Journal of Computer Trends and Technology 68.4 (2020):126-128.

Verma, P., Pal, S., and Om, H. (2019). A comparative analysis on hindi and english extractive text summarization. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):30. Wan, X. (2010).

Widyaningrum, I. ., Wibisono, N. ., & Kusumawati, A. H. . (2020). Effect of extraction method on antimicrobial activity against staphylococcus aureus of tapak liman (elephantopus scaber l.) leaves. International Journal of Health & Medical Sciences, 3(1), 105-110. https://doi.org/10.31295/ijhms.v3n1.181