



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

**Volume 10, Issue 5, May 2022**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.165**



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

# Blog Mining and Emotion Argumentation

Ms. K. Vineela, M. Sruthi, M. Sravya, K. Saraswathi, K. Veena Madhuri

Assistant Professor, Department of IT, Vasireddy Venkatadri Institute of Technology, Nambur, A.P, India

B. Tech Students, Department of IT, Vasireddy Venkatadri Institute of Technology, Nambur, A.P, India

**ABSTRACT:** Blogs are perhaps widely used by the internet users due to their ability to disseminate information and present their ideas on various topics. These blogs have increasingly become an important information source for the users' ideas. So, we are providing a platform to Blogger's and Reader's, where the blogger will continually update new content or information in the Blog and readers will check them frequently and gain knowledge or information from the blogs. These subjective information in blogs helps in understanding a blogger's views and observations about various topics. So, we created a website providing access for users and bloggers, where bloggers have a separate account in the website, they can login and update/write the content and readers can go to the home page directly to view the blogs of bloggers. Here we are creating a Techie Blog website. Blog Mining and Emotion Argumentation provides a capability of processing large amount of text data effectively, it can be a valuable method for gaining insights into a given topic. Due to its capability of processing large amount of text data effectively, the overloaded content is summarized and the user can easily deluge huge amount of information in a shorter span and also allowing quicker search results.

The steps involved here are:

1. Pre-processing of data
2. Thus, pre-processed data is updated by blogger
3. The info updated is summarized (involves text summarization)
4. Readers browse the content and access necessary blog where deluge content is summarized and user gain insights about a blog in shorter span and also having fast search results.

**KEYWORDS:** Blog Mining, Emotion Argumentation, Preprocessing of data, Deluge content is shortened, Text Summarization, Latent Semantic Analysis, Document term matrix, Singular value decomposition. Visualize condensed impact in graphical view.

## I. INTRODUCTION

A blog (a shortened version of "weblog") is an online journal or informational website displaying information in reverse chronological order, with the latest posts appearing first, at the top. It is an informational website, often informal diary-style text entries, and acting as a platform where a writer or a group of writers share their views on an individual subject. It is similar to an online journal where an individual, group, or corporation presents a record of activities, thoughts, or beliefs. In general terminology blog is explained as a regularly updated website or web page, typically run by an individual or small group, that is written in an informal or conversational style, consisting of series of posts where posts are archived, and are usually sorted into categories. It is similar to a newspaper in that it publishes new items on a regular basis and keeps the older ones up to date. Bloggers identify the sentiments, both positive and negative opinions about the topic to understand and present public views in detail. Readers can browse these categories through the blog to read older entries. It does typically involve searching and analysing blogs in order to generate additional insights and acts as an information source for the user's ideas. Blog Mining and Emotion Argumentation provides a capability of processing large amount of text data effectively. This study of blog mining is used to analyse and search the online blog posts relevant contents in a quite simpler fashion.

## II. RELATED WORK

### Aim and Scope –

Blog mining and Emotion Argumentation allows to extract the huge amount of corpus and allows to view it in as condensed state. An individual user diverges respective contents of blog, the encapsulated content allows to digest the huge info in a compact way, consuming minimal time impact. The processing of contents via browsing reduces the actual time consumption to lower. The series of blogs when updated via from the end of the user, which there by provides a summarized view updation to end of database. The main constraints involved are blog corpus source, depictive info of a blog and representative contents, Machine Learning view impact to overcome the complication. Argumentation is the process by which arguments are constructed and handled. Argumentation

constitutes a major component of human intelligence. Argumentation is a collection of propositions, all of which are premises except, at most one, which is a conclusion.

#### **Goals and Objectives –**

The typical goals of the Blog mining and Emotion Argumentation are to collect the blog corpus and design a system for topic identification and other text processing tasks such as text summarization unit. As the amount of text keeps growing, it becomes increasingly difficult for humans to process the deluge of information in the time available. It does result in consumption of huge amount of duration. As a part, the subject propagated within the blogs may not be reachable to end users. Blog Mining overcomes by performing text routing techniques like text summarization. Inclusion of Emotion Argumentation is used for developing, analyzing and categorizing the arguments. The overloaded content which is extracted from huge number of posts result huge time to browse as a process of application of summarization provides a faster pace of search.

#### **Existing System –**

Blogs are often updated webpages that are being important information sites about a specific topic. Food, fashion, and marketing are just a few of the topics covered by various sorts of blogs. Here the source that we have taken is techie blog. Generally, all posts are archived in blog, and are usually sorted into categories. Readers can browse these categories of a blog to read older entries. It does typically involve searching and analyzing blogs in order to generate additional insights and acts as a information source for the user's ideas.

#### **Disadvantages –**

1. The overloaded content which is extracted from huge number of posts provides out low search results consuming huge time.
2. As the amount of text keeps growing, it becomes increasingly difficult for humans to process the deluge of information in the time available. It does result in consumption of huge amount of duration. As a part of low efficiency, the subject propagated within the blogs may not be reachable to end users.

#### **Problem –**

Bloggers are facing their toughest challenge yet and it's called content saturation. There's just too much information to process these days. And when people feel overwhelmed, they react in ways that aren't good for a respective blog. Information overload occurs when a person is exposed to more information than the brain can process at one time. As the amount of on-line text keeps growing, it becomes increasingly difficult for humans to process the deluge of information in the time available. The rapid growth of blog documents in web and categorizing search applications based on topics motivates to develop a system that provides identification of the blog documents.

#### **Proposed Solution –**

Text summarization is the process of creating a short, accurate, and fluent summary of a longer text document. The main purpose of text summarization is to get the most precise and useful information from a large document and eliminate the irrelevant or less important ones. The process brings out information that is crucial, and also ensures that the meaning of the paragraph stays the same. This helps reduce the time to understand large papers like research articles, without skipping any vital information. The main benefits are makes reading easier, saves time, helps to memorize the information easily, boosts the work rate efficiency. This process allows consumption time to be shorter having quicker search results and emphasizing with Machine learning based text processing tasks such as text summarization unit, and allows to Emotional Argumentation via to find the effect of emotions in the generation of conclusions and evaluate the consistency of emotions from a set of premises to its corresponding conclusion.

### **III. PROPOSED ALGORITHM**

The main aim of latent semantic analysis is to create representations of text data in terms of the features and latent features. Latent semantic analysis (LSA) is a mathematical method for computer modelling and simulation of the meaning of words and passages by analysis of representative corpora of natural text. LSA closely approximates many aspects of human language learning and understanding. It supports a variety of applications in information retrieval. The latent semantic analysis consists of two steps:

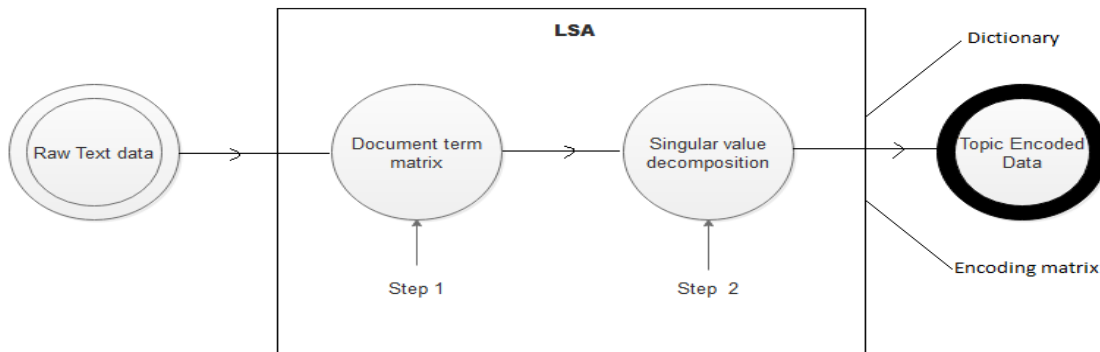


Fig. LSA processing

Step 1: Document term matrix

A document-term matrix is a mathematical matrix that describes the frequency of terms that occur in a collection of documents. In a document-term matrix, rows correspond to documents in the collection and columns correspond to terms. This matrix is a specific instance of a document-feature matrix where "features" may refer to other properties of a document besides terms. They are useful in the field of natural language processing and computational text analysis.

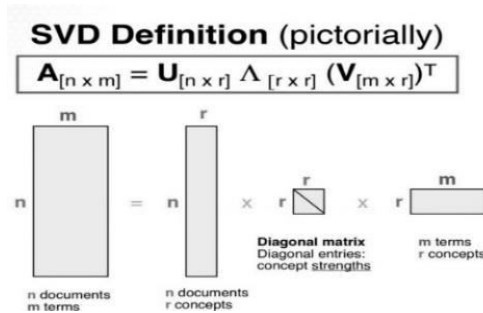
Document Term Matrix

	intelligent	applications	creates	business	processes	bots	are	i	do	intelligence
Doc1	2	1	1	1	1	0	0	0	0	0
Doc2	1	1	0	0	0	1	1	0	0	0
Doc3	0	0	0	1	0	0	0	1	1	1

	brown	dog	fox	lazy	quick	red	slow	the	yellow
"the quick brown fox"	1	0	1	0	1	0	0	1	0
"the slow brown dog"	1	1	0	0	0	0	1	1	0
"the quick red fox"	0	1	0	0	1	1	0	1	0
"the lazy yellow fox"	0	0	1	1	0	0	0	1	1

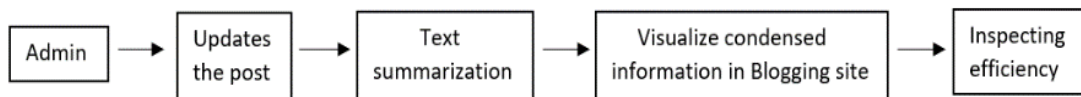
Step 2: Singular Value Decomposition

The Singular Value Decomposition (SVD) of a matrix is a factorization of that matrix into three matrices. It has some interesting algebraic properties and conveys important geometrical and theoretical insights about linear transformations. It also has some important applications in data science. Interpretation of SVD is given by: Deriving mapping between m-dimensional space and r-dimensional singular vector space (rank of input matrix = r). Breaks down the original document into r linearly-independent base vectors or concepts. SVD can semantically cluster words and sentences by finding salient and vectors. The sentence that best represents this pattern will have the largest index value. After you run SVD and get the most salient concepts in the text, we need to select the sentences as summary.

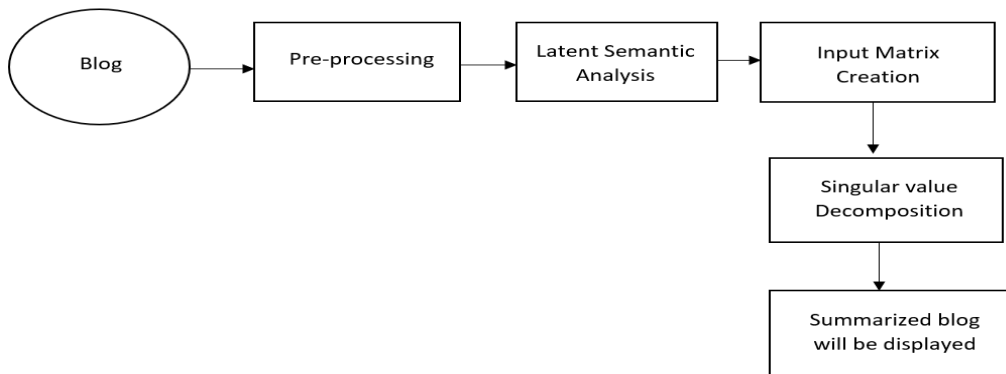


**IV. SYSTEM DESIGN - ARCHITECTURE**

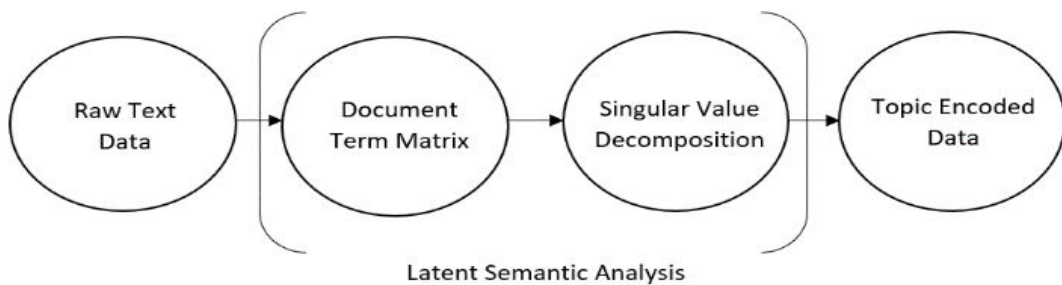
**SYSTEM VIEW**



**TEXT SUMMARIZER**



**LATENT SEMANTIC ANALYSIS**



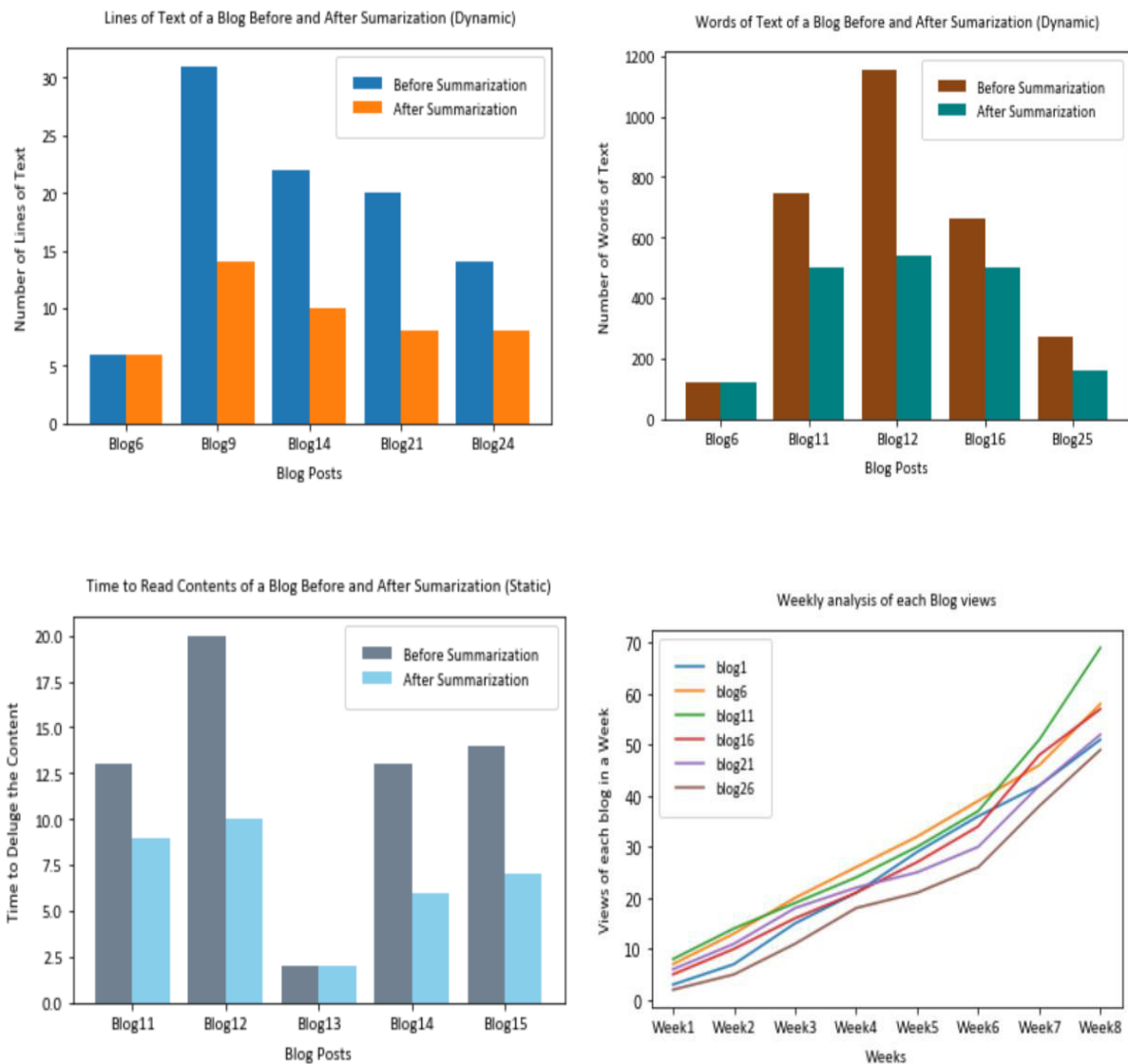
**Description –**

The blog mining involves admin who updates new blog, here by updated content is summarized in short and is stored into the database. Thus, summarized content is here by accessed by users while access of info regard a blog. This a part where the condensed information in blogging site is visualized. The next part is to inspect efficiency by gather of

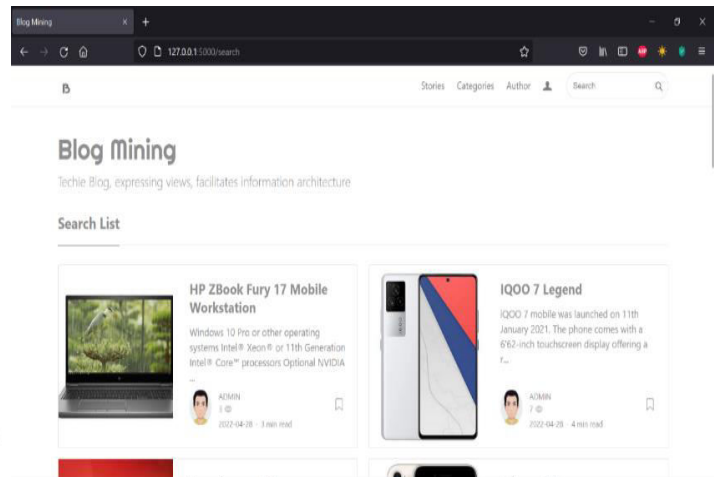
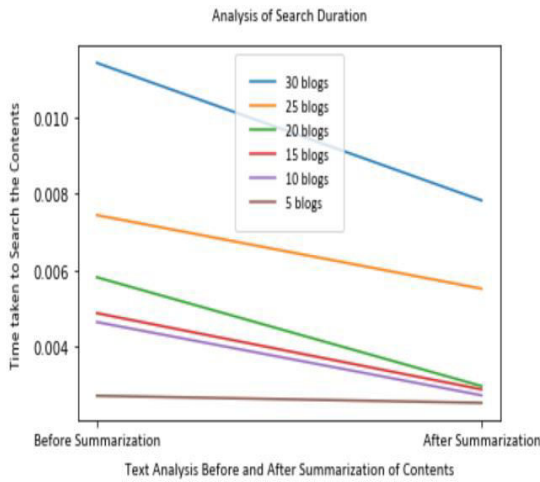
details regard info before summarization and after summarization. Thus, the pre-processed data is allowed to provide it for a graphical representation having a visual impact for comparison of before summarization and after summarization. During this process the tasks involved for text summarizer is blog content is pre-processed applying the algo Latent Semantic Analysis, thus displaying out summarized content of blog. The latent semantic analysis involves typically two steps, first step is document term matrix where the content gathered is represented in a matrix, and then followed by singular value decompose allows to extract the ranked text from the matrix and extracted data which is in encoded matrix format is extracted and pulled out as sentences. Here Argumentation mining occupies a position between natural language processing, argumentation theory and information retrieval. Argumentation mining aims to automatically detect, classify and structure argumentation in text. Argumentation mining focuses on the detection of all the arguments in a text and their relationships with their preceding and following arguments. Argumentation mining does not analyze the validity of the argumentation or its correctness. The aim is to detect those pieces of text which seem to function as argumentative (from a linguistic and semantic point of view) and the relations between them, i.e., their structure. Emotion argumentation means to evaluate the consistency of emotions from a set of premises to its corresponding conclusion.

### V. SIMULATION RESULTS

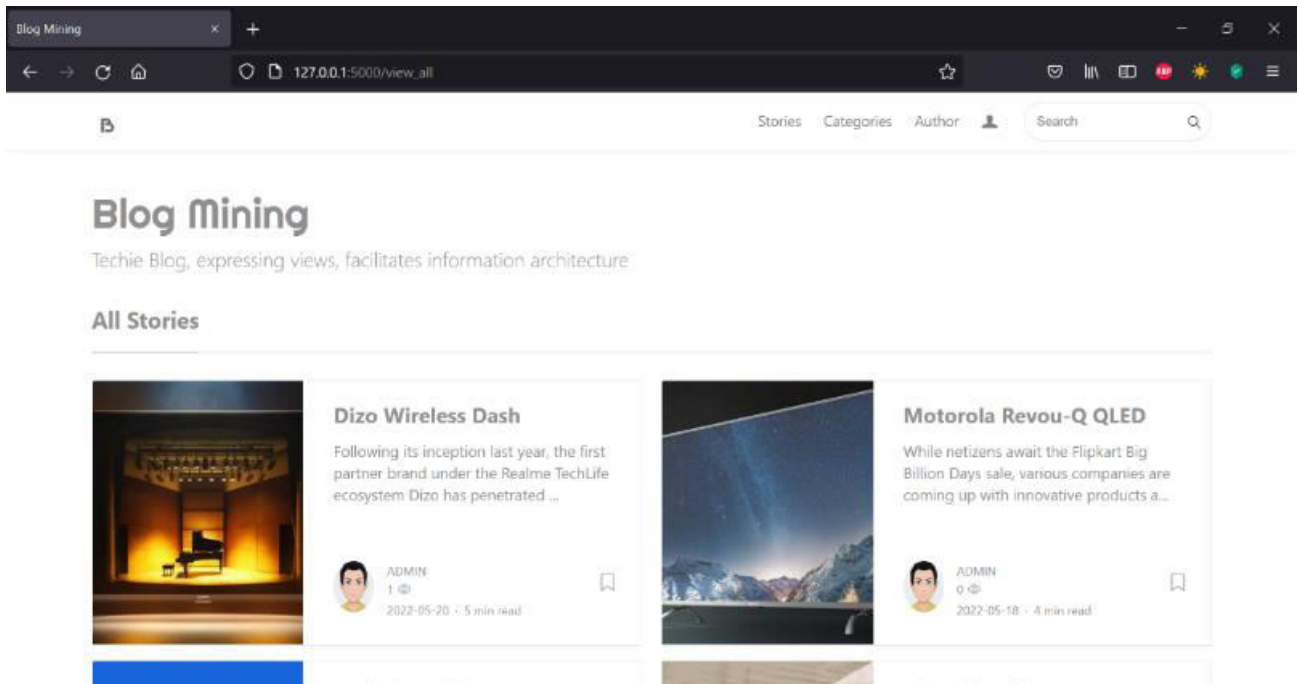
#### ANALYSIS WITH READABILITY, TIME AND VIEWERS –



SEARCH ANALYSIS –



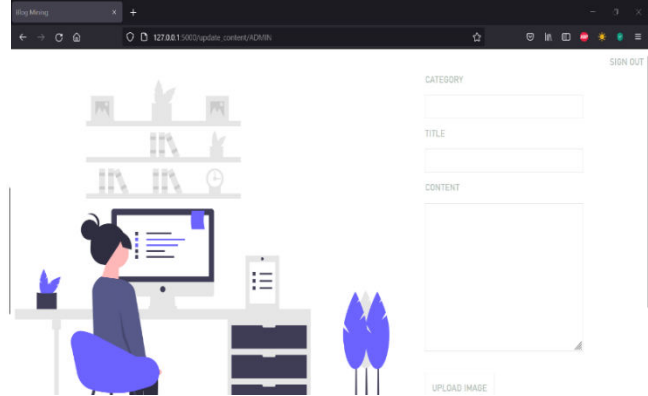
The blog home page –



### The summarized text view in blog – Updation of summarized content by blogger –

Back in September last year, HP announced the world's first 11-inch Windows-powered tablet with a flip camera. Although the device was slated to be up for grabs sometime in December 2021 in the US, the promise could not be delivered at the time. Now, HP's 11-inch tablet has finally hit the shelves in the US. Initially spotted by Gizmochina, the HP 11-inch tablet is currently available on Best Buy with a starting price of \$499 (~Rs 37,264). While the standalone device comes at the above price, customers can also buy the tablet along with a detachable keyboard accessory for a price of \$599 (~Rs 44,732). The device comes in a single Natural Silver color variant and runs Windows 11 S Mode out-of-the-box. Touted as the world's first tablet with a flip camera mechanism, the device comes with a 13MP single rear camera that can flip out from the back to become a high-quality front-facing webcam. The mechanism is similar to the flip camera seen on the Asus ZenFone series, which launched back in 2020. The tablet PC is backed by a 32'2Wh battery that charges via the onboard USB-C port. Plus, the device supports the HP Tilt Pen to draw or take notes and a kickstand for portrait and landscape viewing.

Accessories Laptop Mobile Processor



### VI. CONCLUSION AND FUTURE WORK

The study of blog mining and emotion argumentation is used to analyze and search the online blog posts relevant contents in a quite simpler fashion. The main purpose includes provision of summarize content. As the information reflected is summarized view one can grasp away huge content in shorter duration and also browsing gives faster pace results. Blog mining can be concluded as a view by pulling out search optimization and also allows users to acknowledge deluge information of blogging websites.

In future we can implement with various languages like Hindi. This allows more enrichment for users who prefer native language, the source having first language allows many users to put interest to look through vivid contents of the blogs.

### REFERENCES

- [1] Thomas K. Landauer, Danielle S. McNamara, Simon Dennis, Walter Kintsch, Handbook of Latent Semantic Analysis
- [2] Erkan and D. Radev. 2004. Lexpagerank: Prestige in multi-document text summarization.
- [3] Kazantseva and S. Szpakowicz. 2006. Challenges in evaluating summaries of short stories.
- [4] D'Avanzo and B. Magnini. 2005. A keyphrase-based approach to summarization: The Lake system at DUC 2005. In *Proceedings of the Document Understanding Conference (DUC 2005)*.
- [5] Mehwish Aziz, Muhammad Rafi, Sentence-Based Semantic Similarity measure for Blog Posts, 6th International Conference on Digital Content, Multimedia Technology and its Applications, IDC 2010, Aug 2010, IEEE Seoul.
- [6] Alterman R., "Summarization in the small" in N. Sharkey edition – Advances in cognitive science. Chichester, England, Ellis Horwood. 1986.
- [7] Alterman R., "Text summarisation" in Artificial Intelligence Review. 1990.
- [8] Baxendale, P.B., "Man-made index for technical literature: an experiment", IBM Journal of Research and Development, 2, 4, 1958.
- [9] Aho, A., Chang, S.-F., McKeown K, Radev, D., Smith, J., and Zaman, K. 1997. "Columbia digital news system: An environment for briefing ansearch over multimedia information". In Proceedings of IEEE ADL, Washington, DC.
- [10] Edmundson H.P., "New methods in automatic extracting" in Journal of the ACM, 6, 264-285. 1969.
- [11] Cohen, J.D., "Highlights: Language- and Domain-Independent Automatic Indexing Terms for Abstracting", Journal of the American Society for Information Science.
- [12] Garner, R., Efficient text summarization: costs and benefits, Journal of Educational Research.

### BIOGRAPHY

We **Sruthi Mandalapu, Sravya Mannava, Saru Koduru, Veena Madhuri Konakanchi** students in the Information Technology Department, College of Vasireddy Venkatadri Institute of Technology, JNTUK, guided by **Ms. Vineela Kanakamedala (Assistant Professor)**.





**INNO**  **SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.165**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
**INDIA**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details