

# TEXT SUMMARIZATION

Presented By:

Vedant Landge - 20190802005

Krish Shetty - 20190802129

# TABLE OF CONTENTS

**01**

**INTRODUCTION**

**03**

**RESULTS**

**02**

**TEXT  
SUMMARIZATION**

**04**

**CONCLUSION**

The background is a dark blue gradient. It features several abstract geometric elements: a network of light blue lines with circular nodes at the top center; a large, multi-layered 3D triangle on the left side; and several smaller 3D triangles scattered on the right and bottom. The main title is centered in a large, bold, white sans-serif font.

# PROBLEM STATEMENT


To extract summary of a long text in order to help us  
gain faster insights about a piece of article or even  
research papers

Abstract geometric shapes, including triangles and polygons, in various shades of blue and white, located in the top-left corner of the slide.

**01**

# **INTRODUCTION**

Abstract geometric shapes, including triangles and polygons, in various shades of blue and white, located in the bottom-right corner of the slide.

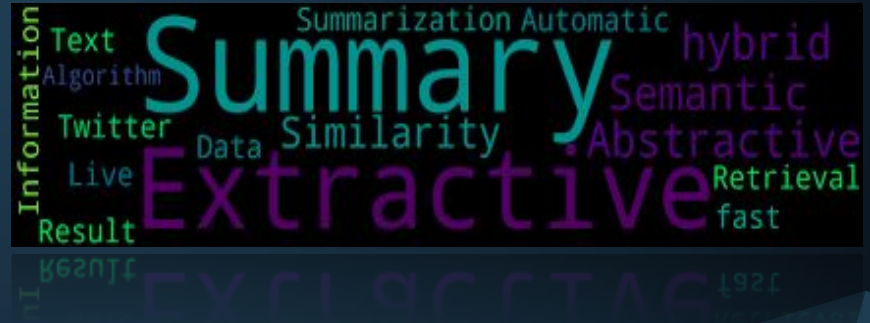
The background is a dark blue gradient. It features several abstract geometric elements: a large, complex, multi-layered triangle in the top-left corner; a network of thin white lines with small circular nodes in the bottom-left corner; and a large, solid triangle in the bottom-right corner, with a smaller, hollow triangle nested inside it. The text is centered in the middle of the image.

“We don’t need a full report. Just give me a  
summary of what you did”

**—SOME CORPORATE SUPERIOR**

# OBJECTIVE

- Amount of textual data is increasing rapidly day-by-day
- Social Media, News articles, emails, etc., generate massive amounts of information
- Data has become the oil of this century



# Total Amount of Digital Data

**4.4 zettabytes**

in 2013

**180 zettabytes**

in 2025.



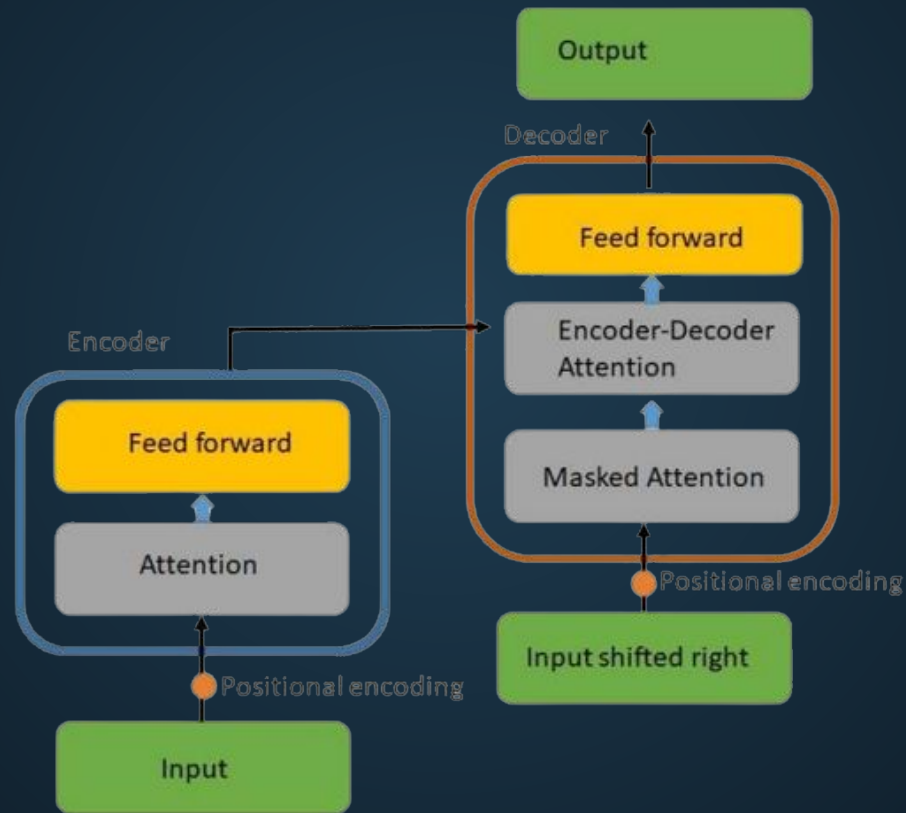
## **02** METHODOLOGY



# TEXT SUMMARIZATION

- Key natural language processing (NLP) task
- Automatically converts a text, or a collection of texts within the same topic, into a concise summary
- Greatly needed to address the ever-growing amount of text data
- Help discover relevant information and consume relevant information faster

# TRANSFORMER ARCHITECTURE



# ALGORITHM

DATA PREPROCESSING  
AND DATASET PIPELINE

POSITIONAL  
ENCODING

FEED FORWARD  
NETWORK

TRAIN THE MODEL  
AND THEN USE IT FOR  
SUMMARIZATION



# DATASET

1	Headline	Short	Source	Time	Publish Date
2	4 ex-bank officials booked for cheating bank of ₹209 crore	The CBI on Saturday booked four former	The New Indian Express	09:25 AM	26-Mar-17
3	Supreme Court to go paperless in 6 months: CJI	Chief Justice JS Khehar has said the Su	Outlook	10:18 PM	25-Mar-17
4	At least 3 killed, 30 injured in blast in Sylhet, Bangladesh	At least three people were killed, inclu	Hindustan Times	11:39 PM	25-Mar-17
5	Why has Reliance been barred from trading in futures?	Mukesh Ambani-led Reliance Industries	Livemint	11:08 PM	25-Mar-17
6	Was stopped from entering my own studio at Times Now: Ar	TV news anchor Arnab Goswami has sa	YouTube	11:24 PM	25-Mar-17
7	New trailer of 'Justice League' released	A new trailer for the upcoming superh	YouTube	09:50 PM	25-Mar-17
8	His touch was not right: Shilpa Shinde on sexual harassment	Television actress Shilpa Shinde, while	The Quint	09:18 PM	25-Mar-17
9	Anti-Romeo squads must not trouble consenting youths: Yog	Uttar Pradesh Chief Minister Yogi Adity	ANI	11:05 PM	25-Mar-17
10	Both Romeo and Juliet are welcome in Delhi: AAP Minister	In an apparent jibe at UP's anti-R	India Today	09:26 AM	26-Mar-17
11	30 blasts occur at ordnance factory in MP, 20 injured	At least 20 people were reportedly inj	Hindustan Times	09:24 PM	25-Mar-17
12	Govt dismisses reports of carving out larger Nagaland state	The Centre on Saturday dismissed rep	PTI	11:06 PM	25-Mar-17
13	White man shouts 'you don't belong here' to a Sikh-American woman was allegedly	A Sikh-American woman was allegedly	Hindustan Times	09:13 PM	25-Mar-17
14	Thousands march in London to protest against Brexit	Thousands of people on Saturday took	Hindustan Times	08:45 PM	25-Mar-17
15	US brewer MillerCoors sues HCL Technologies for ₹650 crore	Indian IT major HCL Technologies is be	Outlook	08:52 PM	25-Mar-17
16	Kangana gifts herself 3-bedroom bungalow on birthday: Rep	As per reports, actress Kangana Ranaut	India Today	10:25 PM	25-Mar-17
17	Flipper taught by Warner got me Warner's wicket: Kulde	Indian spinner Kuldeep Yadav, who re	India Today	08:21 AM	26-Mar-17
18	Underprivileged kids to fill Wankhede during MI's IPL m	As a part of an education initiative, the	Sportskeeda	09:26 AM	26-Mar-17
19	There's clear, irrefutable evidence of EVM manipulation	Aam Aadmi Party (AAP) on Saturday al	Hindustan Times	10:27 PM	25-Mar-17
20	55% power tariff hike in Bihar slashed to 28% after protests	Nitish Kumar-led Bihar government ha	The New Indian Express	11:14 PM	25-Mar-17
21	Don't know about assets of wife: Abdullah in poll affide	As per the affidavit filed by NC Preside	The New Indian Express	10:12 PM	25-Mar-17
22	Indian Army considering to reduce its 'sahayaks' b	The Indian Army, after consultations w	India Today	09:15 PM	25-Mar-17
23	UP cop suspended for consuming tobacco	A police officer in Lucknow, Nagesh Mi	ANI	09:32 PM	25-Mar-17

Abstract geometric shapes, including triangles and polygons, in various shades of blue and white, located in the top-left corner of the slide.


**03**

# **SIMULATION**

Abstract geometric shapes, including triangles and polygons, in various shades of blue and white, located in the bottom-right corner of the slide.



# **04 CONCLUSION**

- 
- In this way, we have implemented the Transformer Architecture to summarize a particular block of text
  - This model, we can use in the future for much wider scale of applications like Meeting Summarization

# LITERATURE SURVEY

1. Vishal Gupta and Gurpreet Singh:

“A Survey of Text Summarization Extractive techniques”.

In this paper, the author describes the extractive summarization methods which comprise two parts Pre-Processing and Processing. In this paper, the pre-processing step is further divided into other subprocesses which are sentence segmentation, stop word removal, and stemming. In the processing step, the weights are given to the features used for extraction of summary from the large document respectively.

2. Saranyamol C S and Sindhu L:

“A Survey on Automatic Text Summarization.”

In this paper, the author describes the various techniques used in automatic text summarization which are extractive text summarization and abstractive text summarization.

3. Rafael Ferreira, Luciano de Souza:

"Assessing sentence scoring techniques for extractive text summarization."

This paper gives a brief description of various features used to perform extractive summarization and it also describes the methods for summary evaluation.



# REFERENCES

1. <https://towardsdatascience.com/transformer-networks-a-mathematical-explanation-why-scaling-the-dot-products-leads-to-more-stable-414f87391500>
2. <https://towardsdatascience.com/understand-text-summarization-and-create-your-own-summarizer-in-python-b26a9f09fc70>
3. <https://blog.knoldus.com/what-are-transformers-in-nlp-and-its-advantages/#:~:text=NLP's%20Transformer%20is%20a%20new,relies%20entirely%20on%20self%20attention.>
4. <https://towardsdatascience.com/transformers-89034557de14>



# GITHUB

<https://github.com/krishshetty23/TextSummarization>

<https://github.com/vedant-landge/Text-Summarization>

The background is a dark blue gradient. In the top-left corner, there is a network of white lines connecting several small blue dots. On the right side, there are several 3D geometric shapes, including a large, dark blue, multi-layered triangular prism and a smaller, dark blue, multi-layered triangular prism. In the bottom-left corner, there is a dark blue, multi-layered triangular prism. In the bottom-right corner, there is a dark blue, multi-layered triangular prism.

**THANK  
YOU!**