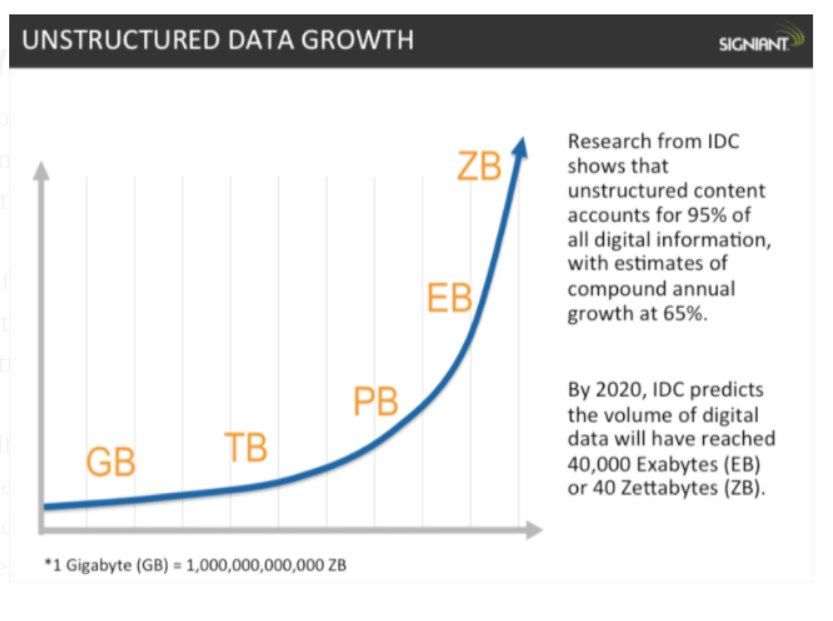
TEXT SUMMARIZATION

# **Why we have chosen the project**

* Reduces reading time.
* Summaries make the selection process of documents easier.
* Improves the effectiveness of indexing.
* Less biased than human summarizers.
* Personalized summaries are useful in question-answering systems.
* Enables commercial abstract services to increase the number of texts they are able to process.



- In fact, the International Data Corporation that the total amount of digital data circulating annually around the world would sprout from 4.4 zettabytes in 2013 to hit 180 zettabytes in 2025. That’s a lot of data!

# **What will be our learnings from the project**

The main learning of the work done in this project are:

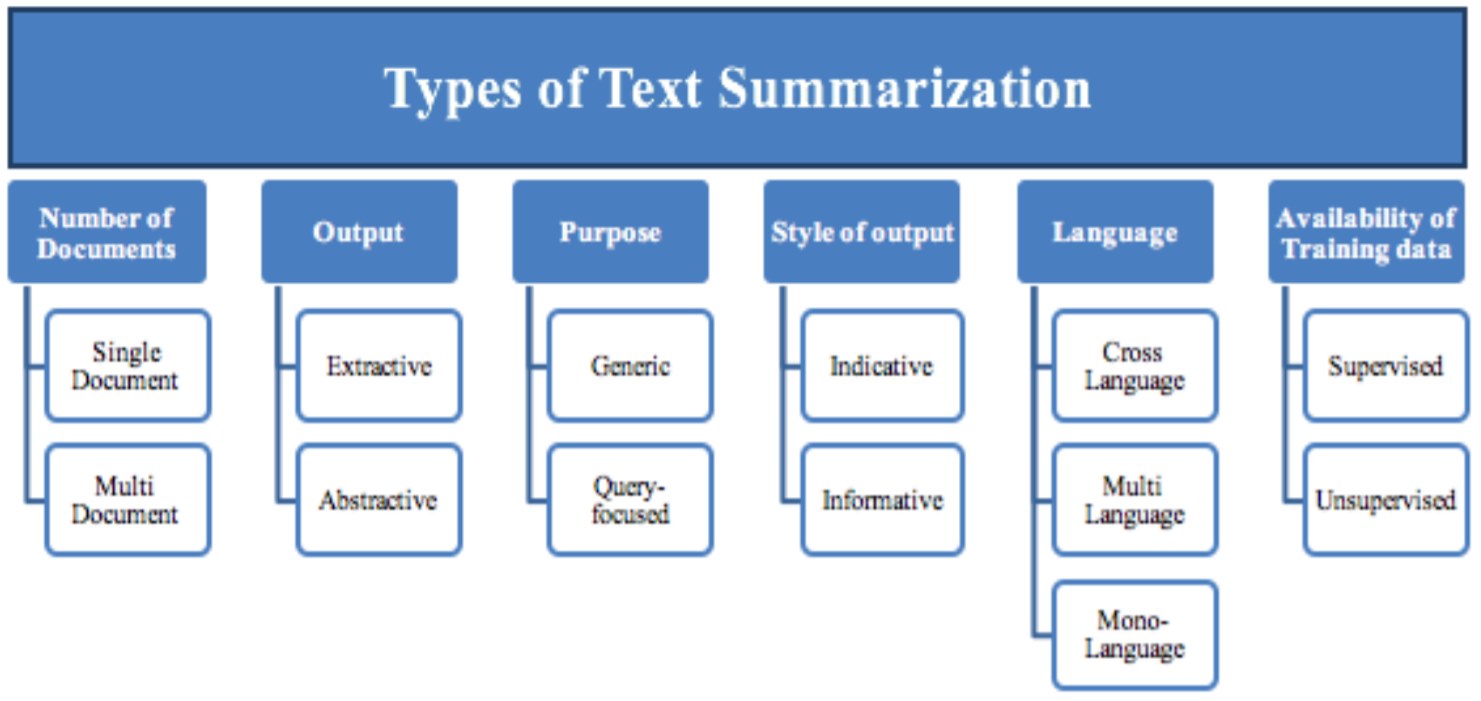
* Extractive algorithm like Fuzzy logic methods.
* Graph Based algorithm like semantic graph method (Textrank, Lexrank )
* Combining different algorithms (i.e., Fuzzy Logic ,Textrank ,Lexrank and LSA)
* Identifying the scope of text summarization algorithms

# **Problem domain**

These are various phases where we faced some problems

* Finding relevant research papers
* Understanding the algorithms that were proposed in those papers
* Implementing the merger system

# **Existing solutions related to our problem**



- Overview of different text summarization methods

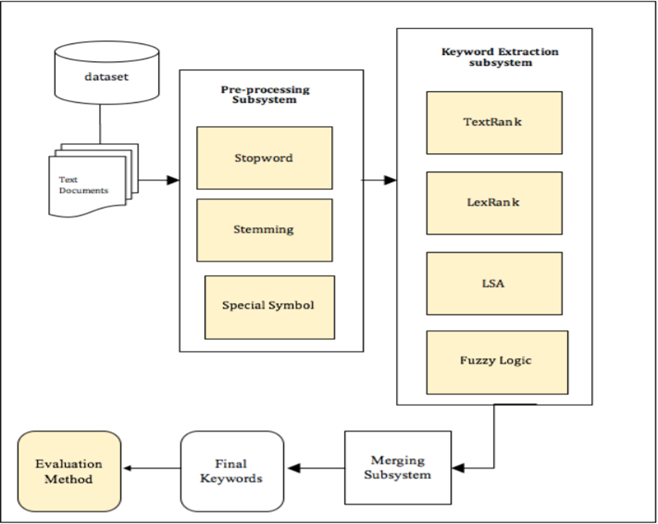
# Proposed Solution

Our Solution includes 4 different methods which are:

* Latent Semantic Analysis
* Fuzzy Logic
* TextRank
* LexRank

Outputs generated by the above are merged by merger system

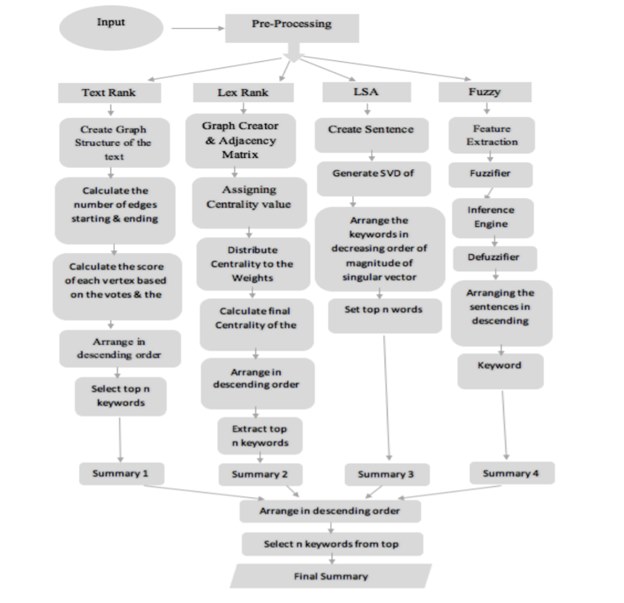
# Architectural Overview



- Architectural Overview

# **Proposed Algorithm**

After running the four algorithms parallely and merging their output, we get the central idea of the document.



- Following figure shows the detailed flow chart of the proposed method containing step-wise procedure of all the four methods.

# **Programming Aspect**

* Every sentence has its scores determining its importance in the document.
* Sentences coming in more than one algorithm are first taken into the final merged document.
* Rest are arranged in descending order of their scores.
* Selection of the number of sentences according to the user's need.

# Future Scope

We can further extend this work by:

* Assigning different weights to each algorithm used.
* Assigning different weights to the features extracted in the process of summary.
* Incorporate more Nature Language-based (NLP) techniques.