

## Abstract

Text summarization refers to the technique of shortening long pieces of text. The intention is to create a coherent and fluent summary having only the main points outlined in the document.

Automatic text summarization is a common problem in machine learning and natural language processing (NLP).

With such a big amount of data circulating in the digital space, there is a need to develop machine-learning algorithms that can automatically shorten longer texts and deliver accurate summaries that can fluently pass the intended messages.

Machine learning models are usually trained to understand documents and distill the useful information before outputting the required summarized texts.

Furthermore, applying text summarization reduces reading time, accelerates the process of researching for information, and increases the amount of information that can fit in an area.

There are two main types of how to summarize text in NLP:

- The extractive text summarization technique involves pulling key phrases from the source document and combining them to make a summary. The extraction is made according to the defined metric without making any changes to the texts.
- The abstraction technique entails paraphrasing and shortening parts of the source document. When abstraction is applied for text summarization in deep learning problems, it can overcome the grammar inconsistencies of the extractive method.

Here, we have tried to analyze the existing methods like unsupervised, supervised, semantic and structure-based.

We have analyzed their scopes and limitations.

We came to the conclusion that the existing methods have certain challenges such as the anaphora problem, cataphora problem, interpretability, long text readability etc.

We proposed certain solutions to overcome those such as enhancing the quality of the dataset by improving the outliers by replacing them with corrected values from human generated inputs.

Text summarization is an interesting machine learning field that is increasingly gaining traction. As research in this area continues, we can expect to see breakthroughs that will assist in fluently and accurately shortening long text documents.