History

Initially pioneered by Hans Peter Luhn in 1950 at IBM.

Developed an algorithm that found “content words” and processed their frequency and significance in a document.

Different approaches based on Single vs Multi-document summarization and Generic summarization vs Query based.

Abstractive Summarization:

More human readable than extractive text summaries.

Abstractive (How is it done?):

Takes a Semantic/Structured approach

Creates a Semantic/Structured representation of the text then uses this to create a summary.

Paraphrases the original text into new sentences.

Abstractive Pre-processing:

Abstractive Scoring:

Example:

Richard drove a car to attend a party in Monaco. In the city, Richard fought with a man named James.

Extractive summarization: Richard drove party Monaco. Richard fought James

Abstractive Summarization: Richard went to Monaco where he fought James.

References:

Andr ́e F.T. Martins, Dipanjan Das. “A Survey on Automatic Text Summarization.” (2007).

Gupta, Som & Gupta, S. K, 2019. Abstractive summarization: An overview of the state of the art. Expert Systems with Applications, 121, pp.49–65.