# Abstract

With the abundance of information on the information on the internet it is difficult for individuals to assess quickly the relevance of a document to them, this is where summaries are beneficial to readers indicating the most salient points of the document and the overall topic it addresses.

To generate quality summaries the initial writer needs to summarise his document, often they find this to be redundant and do not or an unbiased third-party reviewer who understands the topic reads the document and creates a condensed writeup. With the number of documents available and with the number increasing this is infeasible.

Machine learning offers a way to automate summarization of documents quickly and accurately in large corpora of data.

# Introduction

Machine learning is a type of computational algorithmics where the algorithm improves itself as it is run over the corpus of data, “learning”. When implementing Machine learning for automated text summarization is split into main approaches abstractive and extractive, abstractive focused on the individual words and re generates texts and content while extractive instead ranks sentences by prominence.

The results of the summarization are split into two types as well query based where a topical query is given and the machine learning algorithm summarizes based on the query, the alternative a generic summaries that simply condense the most important details of the document.

Research in automated summarization has come to the consensus that summarization models work best on corpora with similar topical content in both results topical similarity helps the algorithm learn.

# Models

Models

# Experiments

Experiments

# Evaluation

Evaluation

# Conclusions

conclusion