Comparative Analysis of Automatic Text Summarization Methods

Ying Luo [yinlu879@student.liu.se]

ABSTRACT

Automatic summarization is the process of shortening a set of data computationally to create a subset (a summary) that represents the most important or relevant information within the original content [1]. With the rise of the information age, there has been an explosion in the amount of text data from various sources. As a result, there is an urgent need for technology to help people obtain information efficiently and quickly from massive texts in various languages, fields, and topics. Therefore, automatic text summarization has become a technology for data mining, information filtering, acquisition, and recommendation.

This paper will implement a basic method based on the concept of Natural Language Processing and compare it with three common text summarization models.

Keywords:

Text Summarization; Extractive Summarization; Abstractive Summarization; TextRank; BERT; Seq2Seq; Pointer-Generator network; ROUGE

INTRODUCTION

In general, there are two approaches for text summarization in NLP: extraction-based summarization and abstraction-based summarization.

Extraction summarization aims to extract the most important sentences directly from the original document to form a summary. It is simple and effective to implement in practical application and performs well in terms of readability because it ensures baseline levels of grammaticality and accuracy.

On the other hand, the purpose of abstraction summarization is to simulate the process of human summarization: understand document → compress information → generate a summary, which means it could generate novel sentences from information extracted from the corpus. It could have some sophisticated abilities such as paraphrasing, generalization, or incorporating real-world knowledge, which is crucial to high-quality summarization and is possible only in an abstractive framework.

BASIC IMPLEMENTATION

In order to understand text summarization better, we first implement a model according to some basic concepts of NLP and compare it with other models. To implement this basic summarization, we adapt unsupervised approaches to evaluate the importance of sentences. After that, according to the compression ratio or the length of the sentences, generate the summary by order. We use three algorithms to evaluate the importance of sentences:

- The frequency of words
- Position in document

- The similarity of sentences

Users can define the weights of these evaluation results to generate the final summary. Code and results can be found here: [https://github.com/kelly-ly/Summarization]

The algorithm based on word frequency

At first, the algorithm would split an input document (stopwords and numbers have been filtered out) into sentences, then compute their TF-IDF, returning a matrix in which each row represents the result of the corresponding sentence. Sum up the values of each row, then normalized them as the result of the evaluation.

The algorithm based on sentence position

Studies have shown that the first sentence in each paragraph can best represent and express the content of the whole paragraph. We compute the importance of sentence position by:

$$Score(S_i) = \frac{n - i + 1}{n}$$

S indicates a sentence in the document, i is the position of S and n is the number of sentences in the document.

The algorithm based on sentence similarity

The algorithm would compute W_{ij}, the cosine similarity for sentence i and sentence j based on the TF-IDF matrix. Then for each sentence, sum its similarity and normalize it as the evaluation result W_i.

$$Similarity = cos(\theta) = \frac{S_i \cdot S_j}{\|S_i\| \|S_j\|} = \frac{\sum\limits_{i=1,j=1}^{n} S_i S_j}{\sqrt{\sum\limits_{i=1}^{n} S_i^2} \sqrt{\sum\limits_{j=1}^{n} S_j^2}}$$

Example

We set the compression ratio as 0.3, and the weights for these three algorithms is even, and the result is shown below:

Input	**Output**
The Chrysler Building, the famous art deco New York skyscraper, will be sold for a small fraction of its previous sales price. The deal, first reported by The Real Deal, was for $150 million, according to a source familiar with the deal. Mubadala, an Abu Dhabi investment fund, purchased 90% of the building for $800 million in 2008. Real estate firm Tishman Speyer had owned the other 10%. The buyer is RFR Holding, a New York real estate company. Officials with Tishman and RFR did not immediately respond to a request for comments. It's unclear when the deal will close. The building sold fairly quickly after being publicly placed on the market only two months ago. The sale was handled by CBRE Group. The incentive to sell the building at such a huge loss was due to the soaring rent the owners pay to Cooper Union, a New York college, for the land under the building. The rent is rising from $7.75 million last year to $32.5 million this year to $41 million in 2028. Meantime, rents in the building itself are not rising nearly that fast. While the building is an iconic landmark in the New York skyline, it is competing against newer office towers with large floor-to-ceiling windows and all the modern amenities. Still the building is among the best known in the city, even to people who have never been to New York. It is famous for its triangle-shaped, vaulted windows worked into the stylized crown, along with its distinctive eagle gargoyles near the top. It has been featured prominently in many films, including Men in Black 3, Spider-Man, Armageddon, Two Weeks Notice and Independence Day. The previous sale took place just before the 2008 financial meltdown led to a plunge in real estate prices. Still there have been a number of high profile skyscrapers purchased for top dollar in recent years, including the Waldorf Astoria hotel, which Chinese firm Anbang Insurance purchased in 2016 for nearly $2 billion, and the Willis Tower in Chicago, which was formerly known as Sears Tower, once the world's tallest. Blackstone Group (BX) bought it for $1.3 billion 2015. The Chrysler Building was the headquarters of the American automaker until 1953, but it was named for and owned by Chrysler chief Walter Chrysler, not the company itself. Walter Chrysler had set out to build the tallest building in the world, a competition at that time with another Manhattan skyscraper under construction at 40 Wall Street at the south end of Manhattan. He kept secret the plans for the spire that would grace the top of the building, building it inside the structure and out of view of the public until 40 Wall Street was complete. Once the competitor could rise no higher, the spire of the Chrysler building was raised into view, giving it the title.	The Chrysler Building, the famous art deco New York skyscraper, will be sold for a small fraction of its previous sales price. The deal, first reported by The Real Deal, was for $150 million, according to a source familiar with the deal. Mubadala, an Abu Dhabi investment fund, purchased 90% of the building for $800 million in 2008. The buyer is RFR Holding, a New York real estate company. The incentive to sell the building at such a huge loss was due to the soaring rent the owners pay to Cooper Union, a New York college, for the land under the building. Still there have been a number of high profile skyscrapers purchased for top dollar in recent years, including the Waldorf Astoria hotel, which Chinese firm Anbang Insurance purchased in 2016 for nearly $2 billion, and the Willis Tower in Chicago, which was formerly known as Sears Tower, once the world's tallest.

Input	Output
In an attempt to scale up its AI portfolio, Apple has acquired Spain-based AI video startup — Vilynx for approximately $50 million. Reported by Bloomberg, the AI startup — Vilynx is headquartered in Barcelona, which is known to build software using computer vision to analyse a video's visual, text, and audio content with the goal of "understanding" what's in the video. This helps it categorising and tagging metadata to the videos, as well as generate automated video previews, and recommend related content to users, according to the company website. Apple told the media that the company typically acquires smaller technology companies from time to time, and with the recent buy, the company could potentially use Vilynx's technology to help improve a variety of apps. According to the media, Siri, search, Photos, and other apps that rely on Apple are possible candidates as are Apple TV, Music, News, to name a few that are going to be revolutionised with Vilynx's technology. With CEO Tim Cook's vision of the potential of augmented reality, the company could also make use of AI-based tools like Vilynx. The purchase will also advance Apple's AI expertise, adding up to 50 engineers and data scientists joining from Vilynx, and the startup is going to become one of Apple's key AI research hubs in Europe, according to the news. Apple has made significant progress in the space of artificial intelligence over the past few months, with this purchase of UK-based Spectral Edge last December, Seattle-based Xnor.ai for $200 million and Voysis and Inductiv to help it improve Siri. With its habit of quietly purchasing smaller companies, Apple is making a mark in the AI space. In 2018, CEO Tim Cook said in an interview that the company had bought 20 companies over six months, while only six were public knowledge.	In an attempt to scale up its AI portfolio, Apple has acquired Spain-based AI video startup — Vilynx for approximately $50 million. Apple told the media that the company typically acquires smaller technology companies from time to time, and with the recent buy, the company could potentially use Vilynx's technology to help improve a variety of apps. The purchase will also advance Apple's AI expertise, adding up to 50 engineers and data scientists joining from Vilynx, and the startup is going to become one of Apple's key AI research hubs in Europe, according to the news.

OTHER MODELS

TextRank

The importance of a sentence is not only reflected in the internal word of the sentence but also the relationship between the sentences in the document: if more support from other sentences, the more important the sentence is.

TextRank [2] is a graph-based ranking model for text processing that can be used to find the most relevant sentences and keywords in the text. It aims to rank the phases in the document and then select K key-phrases to evaluate each sentence. In the end, according to compression ratio to select sentences and sort them as original ordering.

TextRank would create an undirected graph $G(V, E)$, in which vertices V represents sentence set and edges E is set of undirected edges. If there is $e = (V_i, V_j) \in E$, it means sentence V_i and V_j have relevance or similarity, and it is represented by the weight W_{ij}.

$$W_{ij} = \frac{|\{w_k | w_k \in V_i \ \& \ w_k \in V_j\}|}{log|V_i| + log|V_j|}$$

The vertex weight formula is:

$$WS(V_i) = (1-d) + d * \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}} WS(V_j)$$

The sentence weight formula is (S stands for sentence):

$$WS(S_i) = \sqrt{\sum_{V_j \in S_i} (V_j^2)}$$

We use the same documents to test, and the results are as follows(compression ratio = 0.3, number of key-phrases K = 4) [3]:

Input	**Output**
The Chrysler Building, the famous art deco New York skyscraper, will be sold for a small fraction of its previous sales price. The deal, first reported by The Real Deal, was for $150 million, according to a source familiar with the deal. Mubadala, an Abu Dhabi investment fund, purchased 90% of the building for $800 million in 2008. Real estate firm Tishman Speyer had owned the other 10%. The buyer is RFR Holding, a New York real estate company. Officials with Tishman and RFR did not immediately respond to a request for comments. It's unclear when the deal will close. The building sold fairly quickly after being publicly placed on the market only two months ago. The sale was handled by CBRE Group. The incentive to sell the building at such a huge loss was due to the soaring rent the owners pay to Cooper Union, a New York college, for the land under the building. The rent is rising from $7.75 million last year to $32.5 million this year to $41 million in 2028. Meantime, rents in the building itself are not rising nearly that fast. While the building is an iconic landmark in the New York skyline, it is competing against newer office towers with large floor-to-ceiling windows and all the modern amenities. Still the building is among the best known in the city, even to people who have never been to New York. It is famous for its triangle-shaped, vaulted windows worked into the stylized crown, along with its distinctive eagle gargoyles near the top. It has been featured prominently in many films, including Men in Black 3, Spider-Man, Armageddon, Two Weeks Notice and Independence Day. The previous sale took place just before the 2008 financial meltdown led to a plunge in real estate prices. Still there have been a number of high profile skyscrapers purchased for top dollar in recent years, including the Waldorf Astoria hotel, which Chinese firm Anbang Insurance purchased in 2016 for nearly $2 billion, and the Willis Tower in Chicago, which was formerly known as Sears Tower, once the world's tallest. Blackstone Group (BX) bought it for $1.3 billion 2015. The Chrysler Building was the headquarters of the American automaker until 1953, but it was named for and owned by Chrysler chief Walter Chrysler, not the company itself. Walter Chrysler had set out to build the tallest building in the world, a competition at that time with another Manhattan skyscraper under construction at 40 Wall Street at the south end of Manhattan. He kept secret the plans for the spire that would grace the top of the building, building it inside the structure and out of view of the public until 40 Wall Street was complete. Once the competitor could rise no higher, the spire of the Chrysler building was raised into view, giving it the title.	The Chrysler Building, the famous art deco New York skyscraper, will be sold for a small fraction of its previous sales price. The buyer is RFR Holding, a New York real estate company. The incentive to sell the building at such a huge loss was due to the soaring rent the owners pay to Cooper Union, a New York college, for the land under the building. While the building is an iconic landmark in the New York skyline, it is competing against newer office towers with large floor-to-ceiling windows and all the modern amenities. Still the building is among the best known in the city, even to people who have never been to New York. Walter Chrysler had set out to build the tallest building in the world, a competition at that time with another Manhattan skyscraper under construction at 40 Wall Street at the south end of Manhattan.

Input	Output
In an attempt to scale up its AI portfolio, Apple has acquired Spain-based AI video startup — Vilynx for approximately $50 million. Reported by Bloomberg, the AI startup — Vilynx is headquartered in Barcelona, which is known to build software using computer vision to analyse a video's visual, text, and audio content with the goal of "understanding" what's in the video. This helps it categorising and tagging metadata to the videos, as well as generate automated video previews, and recommend related content to users, according to the company website. Apple told the media that the company typically acquires smaller technology companies from time to time, and with the recent buy, the company could potentially use Vilynx's technology to help improve a variety of apps. According to the media, Siri, search, Photos, and other apps that rely on Apple are possible candidates as are Apple TV, Music, News, to name a few that are going to be revolutionised with Vilynx's technology. With CEO Tim Cook's vision of the potential of augmented reality, the company could also make use of AI-based tools like Vilynx. The purchase will also advance Apple's AI expertise, adding up to 50 engineers and data scientists joining from Vilynx, and the startup is going to become one of Apple's key AI research hubs in Europe, according to the news. Apple has made significant progress in the space of artificial intelligence over the past few months, with this purchase of UK-based Spectral Edge last December, Seattle-based Xnor.ai for $200 million and Voysis and Inductiv to help it improve Siri. With its habit of quietly purchasing smaller companies, Apple is making a mark in the AI space. In 2018, CEO Tim Cook said in an interview that the company had bought 20 companies over six months, while only six were public knowledge.	This helps it categorising and tagging metadata to the videos, as well as generate automated video previews, and recommend related content to users, according to the company website. Apple told the media that the company typically acquires smaller technology companies from time to time, and with the recent buy, the company could potentially use Vilynx's technology to help improve a variety of apps. According to the media, Siri, search, Photos, and other apps that rely on Apple are possible candidates as are Apple TV, Music, News, to name a few that are going to be revolutionised with Vilynx's technology.

Bert-extractive-summarizer

Bidirectional Encoder Representations from Transformers (BERT) is a Transformer-based machine learning technique for natural language processing (NLP) pre-training developed by Google [4]. In recent years, the BERT model has been widely used in the field of Natural Language Processing. Unlike other language representation models, BERT is designed to pre-train deep bidirectional representations from an unlabeled text by jointly conditioning on both left and right context in all layers. As a result, the pre-trained BERT model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task-specific architecture modifications [5].

Bert-extractive-summarizer is designed to execute extractive summarization from lectures. It adopts the BERT model for text embeddings and K-Means clustering to identify sentences closest to the centroid for summary selection [6].

The algorithm is as follows:

Step 1: Using Spacy to tokenize the input.

Step 2: Filtering sentences by minimum and maximum length.

Step 3: Modeling by BERT.

Step 4: Using K-Means clustering to compute centroid (K depends on num-sentences or compression ratio).

Step 5: Finding the closest sentence to each centroid.

Step 6: Sorting the selected sentences as the original ordering.

We still use the same documents to test (min_length = 40, max_length = 600, model = bert-large-uncased and ratio = 0.3) [7]:

Input	**Output**
The Chrysler Building, the famous art deco New York skyscraper, will be sold for a small fraction of its previous sales price. The deal, first reported by The Real Deal, was for $150 million, according to a source familiar with the deal. Mubadala, an Abu Dhabi investment fund, purchased 90% of the building for $800 million in 2008. Real estate firm Tishman Speyer had owned the other 10%. The buyer is RFR Holding, a New York real estate company. Officials with Tishman and RFR did not immediately respond to a request for comments. It's unclear when the deal will close. The building sold fairly quickly after being publicly placed on the market only two months ago. The sale was handled by CBRE Group. The incentive to sell the building at such a huge loss was due to the soaring rent the owners pay to Cooper Union, a New York college, for the land under the building. The rent is rising from $7.75 million last year to $32.5 million this year to $41 million in 2028. Meantime, rents in the building itself are not rising nearly that fast. While the building is an iconic landmark in the New York skyline, it is competing against newer office towers with large floor-to-ceiling windows and all the modern amenities. Still the building is among the best known in the city, even to people who have never been to New York. It is famous for its triangle-shaped, vaulted windows worked into the stylized crown, along with its distinctive eagle gargoyles near the top. It has been featured prominently in many films, including Men in Black 3, Spider-Man, Armageddon, Two Weeks Notice and Independence Day. The previous sale took place just before the 2008 financial meltdown led to a plunge in real estate prices. Still there have been a number of high profile skyscrapers purchased for top dollar in recent years, including the Waldorf Astoria hotel, which Chinese firm Anbang Insurance purchased in 2016 for nearly $2 billion, and the Willis Tower in Chicago, which was formerly known as Sears Tower, once the world's tallest. Blackstone Group (BX) bought it for $1.3 billion 2015. The Chrysler Building was the headquarters of the American automaker until 1953, but it was named for and owned by Chrysler chief Walter Chrysler, not the company itself. Walter Chrysler had set out to build the tallest building in the world, a competition at that time with another Manhattan skyscraper under construction at 40 Wall Street at the south end of Manhattan. He kept secret the plans for the spire that would grace the top of the building, building it inside the structure and out of view of the public until 40 Wall Street was complete. Once the competitor could rise no higher, the spire of the Chrysler building was raised into view, giving it the title.	The Chrysler Building, the famous art deco New York skyscraper, will be sold for a small fraction of its previous sales price. The deal, first reported by The Real Deal, was for $150 million, according to a source familiar with the deal. Real estate firm Tishman Speyer had owned the other 10%. The buyer is RFR Holding, a New York real estate company. The building sold fairly quickly after being publicly placed on the market only two months ago. The incentive to sell the building at such a huge loss was due to the soaring rent the owners pay to Cooper Union, a New York college, for the land under the building. The rent is rising from $7.75 million last year to $32.5 million this year to $41 million in 2028.

Input	Output
In an attempt to scale up its AI portfolio, Apple has acquired Spain-based AI video startup — Vilynx for approximately $50 million. Reported by Bloomberg, the AI startup — Vilynx is headquartered in Barcelona, which is known to build software using computer vision to analyse a video's visual, text, and audio content with the goal of "understanding" what's in the video. This helps it categorising and tagging metadata to the videos, as well as generate automated video previews, and recommend related content to users, according to the company website. Apple told the media that the company typically acquires smaller technology companies from time to time, and with the recent buy, the company could potentially use Vilynx's technology to help improve a variety of apps. According to the media, Siri, search, Photos, and other apps that rely on Apple are possible candidates as are Apple TV, Music, News, to name a few that are going to be revolutionised with Vilynx's technology. With CEO Tim Cook's vision of the potential of augmented reality, the company could also make use of AI-based tools like Vilynx. The purchase will also advance Apple's AI expertise, adding up to 50 engineers and data scientists joining from Vilynx, and the startup is going to become one of Apple's key AI research hubs in Europe, according to the news. Apple has made significant progress in the space of artificial intelligence over the past few months, with this purchase of UK-based Spectral Edge last December, Seattle-based Xnor.ai for $200 million and Voysis and Inductiv to help it improve Siri. With its habit of quietly purchasing smaller companies, Apple is making a mark in the AI space. In 2018, CEO Tim Cook said in an interview that the company had bought 20 companies over six months, while only six were public knowledge.	In an attempt to scale up its AI portfolio, Apple has acquired Spain-based AI video startup — Vilynx for approximately $50 million. In an attempt to scale up its AI portfolio, Apple has acquired Spain-based AI video startup — Vilynx for approximately $50 million. Apple told the media that the company typically acquires smaller technology companies from time to time, and with the recent buy, the company could potentially use Vilynx's technology to help improve a variety of apps. According to the media, Siri, search, Photos, and other apps that rely on Apple are possible candidates as are Apple TV, Music, News, to name a few that are going to be revolutionised with Vilynx's technology. In 2018, CEO Tim Cook said in an interview that the company had bought 20 companies over six months, while only six were public knowledge.

Pointer-Generator Networks

Comparing with the extractive method, the abstractive method is another way of automatic text summarization. It may generate novel words and phrases not featured in the source text – as a human-written abstract usually does. Sequence-to-sequence is a family of machine learning approaches used for language processing and is commonly used for text summarization. However, the baseline sequence-to-sequence model with attention has several shortcomings, such as inaccurately reproducing factual details, an inability to deal with out-of-vocabulary (OOV) words, and repeating themselves, especially for multi-sentence summaries. PointerGenerator Networks [8] is designed for addressing these issues in the context of multi-sentence summaries, and improvements are made in the following aspects:

- Using a hybrid pointer-generator network that can copy words from the source text via pointing, which aids accurate reproduction of information while retaining the ability to produce novel words through the generator.
- Using coverage to keep track of what has been summarized, which discourages repetition.

Baseline sequence-to-sequence attentional model:

The model structure of baseline sequence-to-sequence is a classical Encoder-Decoder model. It firstly uses a single-layer bidirectional LSTM to encode the original text into a two-dimensional vector. This two-dimensional vector is also known as the hidden state h_i, then uses a single-layer unidirectional LSTM to decode these hidden states into intelligible language. On each step t, the decoder receives the word embedding of the previous word e_i (the previous word of the reference summary during training; the previous word emitted by the decoder at test time) and get decoder state s_t.

$$e_i^t = \nu^T tanh(W_h h_i + W_s s_t + b_{attn}) \qquad \nu, W_h, W_s \textit{ and } b_{attn} \textit{ are learnable parameters}$$

The attention distribution a^t:

$$a^t = softmax(e^t)$$

The attention distribution can be viewed as a probability distribution over the source words that tells the decoder where to look to produce the next word. It would be used to produce a weighted sum of the encoder hidden states, known as the context vector h_t^* (can be seen as a fixed-size representation of what has been read from the original text on the current step):

$$h_t^* = \sum_i a_i^t h_i$$

Computing the vocabulary distribution P_{vocab} (a probability distribution over all words in the vocabulary):

$$P_{vocab} = softmax(V'(V[s_t, h_t^*] + b) + b') \qquad V, V', b \textit{ and } b' \textit{ are learnable parameters}$$

Final distribution from which to predict words w:

$$P(w) = P_{vocab}(w)$$

The loss for timestep t is the negative log-likelihood of the target word w_t^*:

$$loss_t = -logP(w_t^*)$$

The loss for the whole sequence is:

$$loss = \frac{1}{T}\sum_{t=0}^{T} loss_t$$

Pointer-generator network

The pointer-generator network combines the baseline sequence-to-sequence model and a pointer network, so it is able to generate words from a fixed vocabulary and copy words by pointing. It introduces a generation probability:

$$p_{gen} = \sigma(w_{h^*}^T h_t^* + w_s^T s_t + w_x^T x_t + b_{ptr})$$

Where w_{h^*}, w_s, w_x *and scalar* b_{ptr} are learnable parameters and σ is the sigmoid function. $p_{gen} \in [0, 1]$ for timestep t is a soft switch to choose between a generated word from the vocabulary by sampling from P_{vocab} and a copied word from the input sequence by sampling from the attention distribution a^t. Therefore, we can get an extended vocabulary for each document by union the vocabulary and words in the source document. The probability distribution over the extended vocabulary:

$$P(w) = p_{gen} P_{vocab}(w) + (1 - p_{gen}) \sum_{i:w_i = w} a_i^t$$

If w is an out-of-vocabulary (OOV) word, $P_{vocab} = 0$; if w does not in the source document, $\sum_{i:w_i=w} a_i^t = 0$.

Coverage mechanism

Another shortcoming of the sequence-to-sequence model is repetition, which is especially pronounced for multi-sentence text. To solve this problem, it adapts the coverage model. The coverage mechanism uses a vector c^t (the sum of attention distributions over all previous decoder timesteps) which is an unnormalized distribution over the source document words representing the degree of coverage that received words from the attention so far.

$$c^t = \sum_{t'=0}^{t-1} a^{t'}$$

The coverage vector is used as the fourth input to the attention mechanism, so the word embedding of the previous word e_i on step t should be changed as:

$$e_i^t = \nu^T tanh(W_h h_i + W_s s_t + w_c c_i^t + b_{attn}) \qquad w_c \ is \ a \ learnable \ parameter \ vector$$

And compute coverage vector as well:

$$covloss_t = \sum_i min(a_i^t, c_i^t)$$

The loss for timestep t would be:

$$loss_t = -logP(w_t^*) + \lambda \cdot covloss_t = -logP(w_t^*) + \lambda \sum_i min(a_i^t, c_i^t)$$

We use the pointer-Generator Networks model (baseline sequence-to-sequence attentional model, pointer-generator network and coverage mechanism, respectively) to execute abstraction-based summarization on the same documents [9]:

Input	The Chrysler Building, the famous art deco New York skyscraper, will be sold for a small fraction of its previous sales price. The deal, first reported by The Real Deal, was for $150 million, according to a source familiar with the deal. Mubadala, an Abu Dhabi investment fund, purchased 90% of the building for $800 million in 2008. Real estate firm Tishman Speyer had owned the other 10%. The buyer is RFR Holding, a New York real estate company. Officials with Tishman and RFR did not immediately respond to a request for comments. It's unclear when the deal will close. The building sold fairly quickly after being publicly placed on the market only two months ago. The sale was handled by CBRE Group. The incentive to sell the building at such a huge loss was due to the soaring rent the owners pay to Cooper Union, a New York college, for the land under the building. The rent is rising from $7.75 million last year to $32.5 million this year to $41 million in 2028. Meantime, rents in the building itself are not rising nearly that fast. While the building is an iconic landmark in the New York skyline, it is competing against newer office towers with large floor-to-ceiling windows and all the modern amenities. Still the building is among the best known in the city, even to people who have never been to New York. It is famous for its triangle-shaped, vaulted windows worked into the stylized crown, along with its distinctive eagle gargoyles near the top. It has been featured prominently in many films, including Men in Black 3, Spider-Man, Armageddon, Two Weeks Notice and Independence Day. The previous sale took place just before the 2008 financial meltdown led to a plunge in real estate prices. Still there have been a number of high profile skyscrapers purchased for top dollar in recent years, including the Waldorf Astoria hotel, which Chinese firm Anbang Insurance purchased in 2016 for nearly $2 billion, and the Willis Tower in Chicago, which was formerly known as Sears Tower, once the world's tallest. Blackstone Group (BX) bought it for $1.3 billion 2015. The Chrysler Building was the headquarters of the American automaker until 1953, but it was named for and owned by Chrysler chief Walter Chrysler, not the company itself. Walter Chrysler had set out to build the tallest building in the world, a competition at that time with another Manhattan skyscraper under construction at 40 Wall Street at the south end of Manhattan. He kept secret the plans for the spire that would grace the top of the building, building it inside the structure and out of view of the public until 40 Wall Street was complete. Once the competitor could rise no higher, the spire of the Chrysler building was raised into view, giving it the title.
Baseline Seq2Seq + Attention	It is the brainchild of the owner of the company. It will be sold at the same time of the $ [UNK] million in 2012. It will be sold at the same time of the $ [UNK] million in 2012.
Pointer-Gen	Chrysler purchased 90% of the building for $800 million in 2008. The deal, first reported by the real deal, was for $150 million. Mubadala, an Abu Dhabi investment fund, purchased 90% of the building for $800 million in 2008.
Pointer-Gen + Coverage	The famous art deco new york skyscraper will be sold for a small fraction of its previous sales price. The deal, first reported by the Real Deal, was for $150 million. Mubadala, an Abu Dhabi investment fund, purchased 90% of the building for $800 million in 2008. the rent is rising from $7.75 million last year to $41 million in 2028.

Input	In an attempt to scale up its AI portfolio, Apple has acquired Spain-based AI video startup — Vilynx for approximately $50 million. Reported by Bloomberg, the AI startup — Vilynx is headquartered in Barcelona, which is known to build software using computer vision to analyse a video's visual, text, and audio content with the goal of "understanding" what's in the video. This helps it categorising and tagging metadata to the videos, as well as generate automated video previews, and recommend related content to users, according to the company website. Apple told the media that the company typically acquires smaller technology companies from time to time, and with the recent buy, the company could potentially use Vilynx's technology to help improve a variety of apps. According to the media, Siri, search, Photos, and other apps that rely on Apple are possible candidates as are Apple TV, Music, News, to name a few that are going to be revolutionised with Vilynx's technology. With CEO Tim Cook's vision of the potential of augmented reality, the company could also make use of AI-based tools like Vilynx. The purchase will also advance Apple's AI expertise, adding up to 50 engineers and data scientists joining from Vilynx, and the startup is going to become one of Apple's key AI research hubs in Europe, according to the news. Apple has made significant progress in the space of artificial intelligence over the past few months, with this purchase of UK-based Spectral Edge last December, Seattle-based Xnor.ai for $200 million and Voysis and Inductiv to help it improve Siri. With its habit of quietly purchasing smaller companies, Apple is making a mark in the AI space. In 2018, CEO Tim Cook said in an interview that the company had bought 20 companies over six months, while only six were public knowledge.
Baseline Seq2Seq + Attention	The company is based on the concept of the company's website. it is based on the company's website, which is based on its website. It is based on the company's website, which is based on its website.
Pointer-Gen	The AI startup — Vilynx is headquartered in Barcelona, which is known to build software using computer vision. This helps it categorising and tagging metadata to the videos, as well as generate automated video previews. This helps it categorising and tagging metadata to the videos, as well as generate automated video previews. This helps it categorising and tagging metadata to the videos, as well as generate automated video previews.
Pointer-Gen + Coverage	The AI startup — Vilynx is headquartered in Barcelona, which is known to build software using computer vision to analyze a video's visual, text. This helps it categorizing and tagging metadata to the videos, as well as generate automated video previews. Apple has made significant progress in the space of artificial intelligence over the past few months. with its habit of quietly purchasing smaller companies, Apple is making a mark in the ai space.

EVALUATION

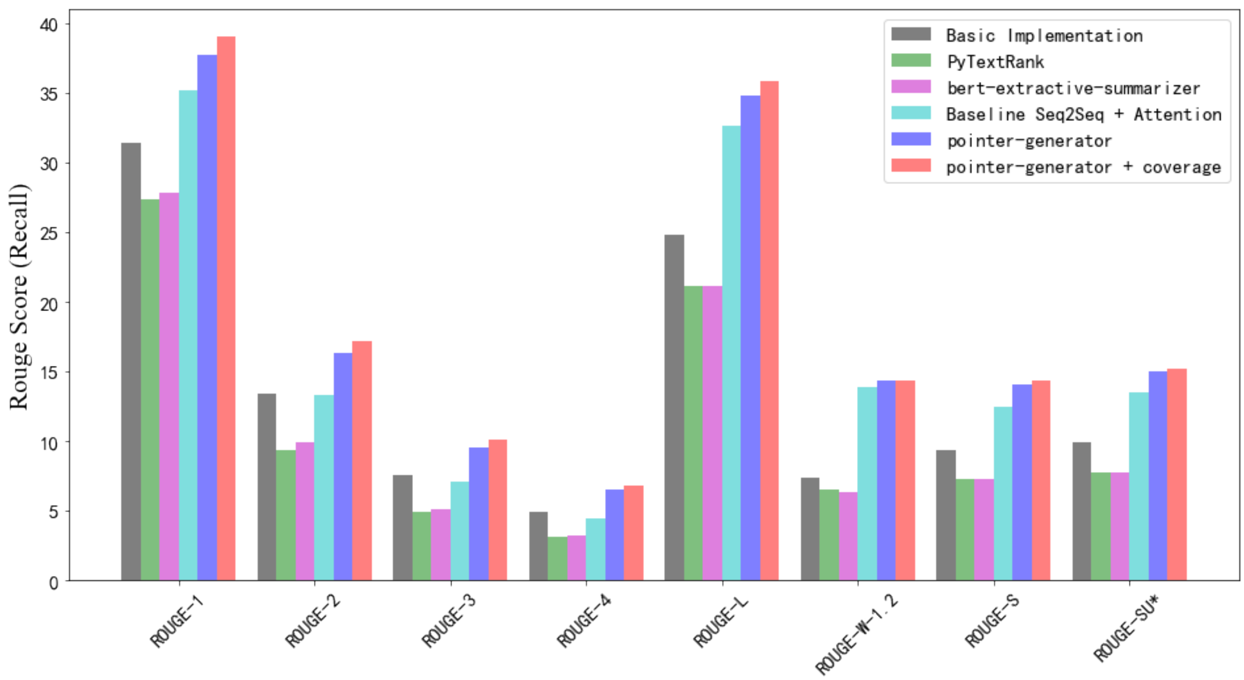
In order to compare and evaluate the performance of summarization models, the most intuitive approach is assessing them manually. The assessment generally considers the following two aspects:

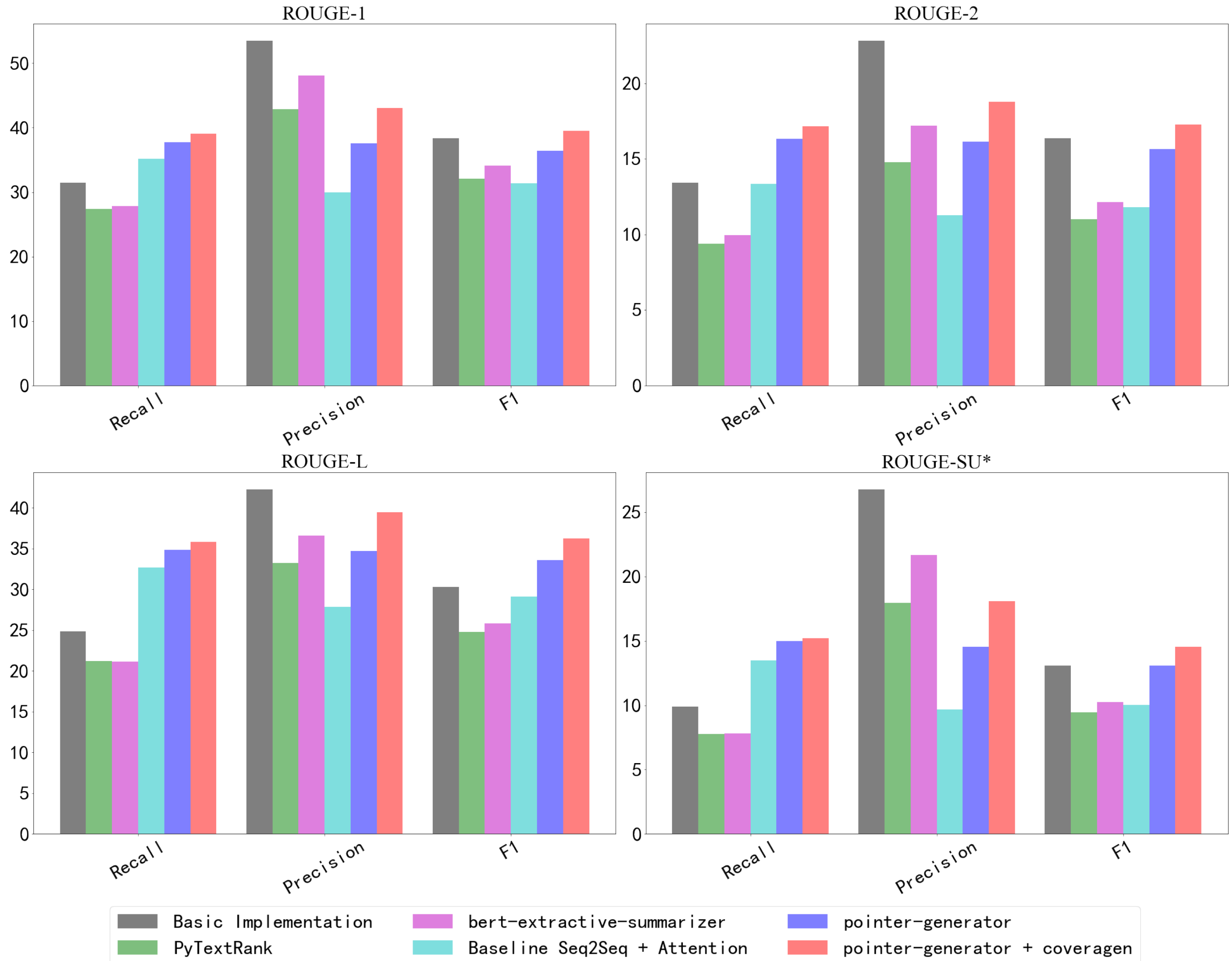
- Whether the generated summary maintains the most important part of the original text.
- How good the readability of summary is, based on grammar and coherence.

However, it is expensive and subjective and is not suitable for evaluating large-scale automatic text summarization. Therefore, automatic evaluation measures for summarization have been proposed. ROUGE [10] is one of the first and most widely used metrics in summarization evaluation.

For the experiment, we use each model to generate summaries and evaluate their performance with respect to the gold human summary. First, we use the pyrouge package to compute the ROUGE scores. The following table shows ROUGE recall results on CNN test set (ROUGE-1, ROUGE-2 and ROUGE-3 are shorthands for unigram and bigram overlap; ROUGE-L is the longest common subsequence; ROUGE-SU means skip-bigram and unigram co-occurrence statistics). All the ROUGE scores have a 95% confidence with at most ± 0.25 interval.

MODEL	**ROUGE-1**	**ROUGE-2**	**ROUGE-3**	**ROUGE-L**	**ROUGE-SU***
Basic Implementation	31.473	13.426	7.596	24.827	9.889
TextRank	26.773	9.263	4.844	20.634	7.467
bert-extractive-summarizer	27.869	9.931	5.144	21.117	7.807
Baseline Seq2Seq + Attention	35.160	13.327	7.134	32.681	13.495
pointer-generator	37.758	16.312	9.581	34.862	14.976
pointer-generator + coverage	**39.051**	**17.156**	**10.079**	**35.820**	**15.212**





From the results, we can find that for the extraction summarization approach, our Basic Implementation performance better than the TextRank and the bert-extractive-summarizer. However, the results of the Pointer-Generator Networks are much better than any extraction summarization approach we tried in this experiment, especially the one adapting coverage mechanism.

Note:

Due to the lack of GPU resources and some optimization parameters, our experimental result may not fully demonstrate the performance of each method. On the other hand, the data set used affects the quality of the summary obtained. For example, the extraction approach might neglect or bias some opinions for those articles with opposing opinions. While ROUGE scores have been shown to often correlate quite well with human judgments [11], they can not be used as the only indicator to evaluate how good a model is.

CONCLUSION

Extraction-based summarization and Abstraction-based summarization have their own advantages and disadvantages. Normally, the extractive approach is simple, efficient, and flexible, and we can set the number of sentences or compression ratio to control the length of the summary. Comparatively, abstraction-based summarization is more closely emulates human summarization, and it has been proved that it performs better than extractive methods in many cases. Nonetheless, since it still has many shortcomings, such as:

- It requires more computing power and time to train and evaluate by deep learning or complex algorithms.

- It may lead to some problems, such as grammar problems, understanding errors, etc.
- Its performance highly depends on the amount and type of training articles.
- It is hard to adjust the length of the final result.

These drawbacks need to be investigated and improved in the future.

We can find from the experimental results that the performance of the abstractive approach is significantly better than the extractive approach. However, we need to specify: there are many methods of automatic text summarization. This paper only discusses three methods that can not determine which platform is absolutely better than the other. In addition, because of extractive methods' simple and effective implementation, the extractive method still plays an important role in many application scenarios.

REFERENCES

1. Automatic summarization. Wikipedia. 2021. https://en.wikipedia.org/w/index.php?title=Automatic_summarization&oldid=1029858948. Accessed 23 Jun 2021.

2. Mihalcea R, Tarau P. TextRank: Bringing order into texts. In: Proceedings of EMNLP-04and the 2004 conference on empirical methods in natural language processing. Barcelonaand Spain; 2004.

3. Nathan P. PyTextRank, a Python implementation of TextRank for phrase extraction and summarization of text documents. 2016. doi:10.5281/zenodo.4602393.

4. BERT (language model). Wikipedia. 2021. https://en.wikipedia.org/w/index.php?title=BERT_(language_model)&oldid=1029082047. Accessed 23 Jun 2021.

5. Devlin J, Chang M-W, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. 2019. http://arxiv.org/abs/1810.04805.

6. Miller D. Leveraging bert for extractive text summarization on lectures. 2019. http://arxiv.org/abs/1906.04165.

7. Miller D. Dmmiller612/bert-extractive-summarizer. 2021. https://github.com/dmmiller612/bert-extractive-summarizer. Accessed 23 Jun 2021.

8. See A, Liu PJ, Manning CD. Get to the point: Summarization with pointer-generator networks. 2017. http://arxiv.org/abs/1704.04368.

9. See A. Abisee/pointer-generator. 2021. https://github.com/abisee/pointer-generator. Accessed 23 Jun 2021.

10. Cohan A, Goharian N. Revisiting summarization evaluation for scientific articles. 2016. http://arxiv.org/abs/1604.00400.

11. Nenkova A, Passonneau R, McKeown K. The pyramid method: Incorporating human content selection variation in summarization evaluation. TSLP. 2007;4.

12. Chen Y, Ma Y, Mao X, Li Q. Multi-Task Learning for Abstractive and Extractive Summarization. Data Science and Engineering. 2019;4:14–23. doi:10.1007/s41019-019-0087-7.