

Research Article

Encoder-Decoder Transformers for Textual Summaries on Social Media Content

Afrodite Papagiannopoulou* , Chrissanthi Angeli 

School of Engineering, University of West Attica, Athens, Greece

Abstract

Social media has a leading role to our lives due to radical upgrade of internet and smart technology. It is the primary way of informing, advertising, exchanging opinions and expressing feelings. Posts and comments under each post shape public opinion on different but important issues making social media's role in public life crucial. It has been observed that people's opinions expressed through social networks are more direct and representative than those expressed in face-to-face communication. Data shared on social media is a cornerstone of research because patterns of social behavior can be extracted that can be used for government, social, and business decisions. When an event breaks out, social networks are flooded with posts and comments, which are almost impossible for someone to read all of them. A system that would generate summarization of social media contents is necessary. Recent years have shown that abstract summarization combined with transfer learning and transformers has achieved excellent results in the field of text summarization, producing more human-like summaries. In this paper, a presentation of text summarization methods is first presented, as well as a review of text summarization systems. Finally, a system based on the pre-trained T5 model is described to generate summaries from user comments on social media.

Keywords

Social Media, Text Summarization, Transformers, Abstractive Summarization

1. Introduction

The internet and social media are flooded these days with an abundance of information which makes it almost impossible for users to read it all. The role of social media is becoming increasingly important especially in public life as users' posts and comments shape public opinion on a variety of important issues concerning politics, economy and society. The role of social media data is of paramount importance as it extracts patterns of social behavior useful in making social, business and government decisions. People tend to express themselves more openly and freely feeling safe in non-direct face to face conversation with strangers. This is due to the

anonymity provided by the internet, the implicit trust in the privacy of communication and the absence of racial stereotypes.

Social media posts are factual. After an event breaks out, social media is flooded with posts related to that event. These posts are in the form of articles, discussions and conversations, the reading of which is a time-consuming and difficult process. Summarizing them afterwards is necessary to retrieve useful knowledge in a reasonable amount of time. Equally important is summarizing the comments of social media users. User comments express the public opinion upon a particular event.

*Corresponding author: apapagiannop@uniwa.gr (Afrodite Papagiannopoulou)

Received: 9 July 2024; **Accepted:** 1 August 2024; **Published:** 15 August 2024



Copyright: © The Author(s), 2024. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Summarizing means converting the content of a long text into a smaller one, preserving the meaning of the original text. There are several ways to write an abstract. It can be done either manually which is called “manual summarization” and it is time consuming or using algorithms and artificial intelligence techniques. The latter is called “Automatic Text Summarization” and is an area that has attracted the interest of researchers especially in recent years [1].

The origins of automated text summarization date back to 1958 [2]. Studies since then have focused on summarizing official documents such as books, journals, scientific articles and technical reports. In these research works, linguistic, heuristic and statistical techniques have been used to create summaries. In modern everyday life, the Internet has now definitively replaced the classical ways of information in the field of news, political and economic developments, advertising and the exchange of opinions, turning the interest of scientists to summary generation systems of web posts, microblogs and social networks. Many algorithms have been implemented to generate text summaries. Two main classes of algorithms can be distinguished: a) Pre-neural, which do not use neural networks; instead they use linguistic and statistical methods to create summaries, b) Deep Learning or Machine Learning techniques that use neural networks. The latter have been successfully applied to various NLP tasks, yielded excellent results, and have been extensively used in recent years [3]. Various deep learning models have been used, mainly based on recurrent neural networks (RNN) and convolutional neural networks (CNN). These models have proven extremely successful in predicting complex relationships that simple structured or semantic approaches cannot do alone [4]. In recent years, the use of Transfer learning and Transformer models has gained popularity mainly in the field of Natural Language Understanding, Processing and Generation [5].

Pre-neural approaches: In the first works, the techniques used to create summaries were related to probabilistic and optimization methods. Twitter was a primary source of receiving posts as input data for the models. In early works [6, 7] the summary is generated by applying a phrase boosting (PR) algorithm and adding the TF-IDF technique to it. The models [8-10] take as input real-time user interaction in the social media stream and provide a multimedia representation as output, considering minimal linguistic information. The model [11] aims to generate a summary from tweets about sports topics. The research is about finding events in real time and includes two stages: 1) Application of the modified Hidden Markov Model by segmenting events based on time into “sub-events” of different importance. 2) Selection of tweets that can provide information about the section that was considered most important. Finally, model [12] produces an information-rich summary based on two different techniques: a) the Decomposition Topic Model (DTM) and b) the Gaussian Decomposition Topic Model (GDTM). These two models exploit the temporal correlation between tweets under predicted conditions.

Neural and Transformer-Based Approaches: The development of Artificial Intelligence and Neural Networks also helped in the further development of NLP. The reasons were twofold: a) the ability to support massive amounts of information on the internet and b) the powerful computing power available today. In the field of summarization research has been done using mainly recurrent neural networks (RNN) and long-term memory (LSTM) RNNs, as well as Convolutional Neural Networks (CNN) [13]. Abstract generation of social media summaries has been done with RNNs based on either Attentional Encoder-Decoder [14], or an attentional awareness mechanism [15] to filter useful information and deal with the specificities of social media content. In [16] an RNN combining sequence-to-sequence and the attention approach is used. The encoder layer is enhanced by an LSTM. The decoder is enhanced with the attention level for more direct access to the input sequence and production of the summaries.

There is a radical change in deep learning applications with the emergence of transformers and the attention mechanism [5]. The idea is adopted by many researchers finding particular resonance in NLP work [17]. Initially transformer models were applied to summarize texts such as articles, books and official texts. A two-stage transformer-based approach to generate abstractive summaries of Chinese articles is proposed [18]. The model produces fluent and variable-length abstracts to meet user requirements. Initially, a pre-trained BERT model and a bidirectional LSTM are used to segment the input text. With the help of mining based BERTSUM model the most important information of the segments is extracted. The model is then trained in two stages based on the transforms. The training process starts with the document Transformer in the second stage. The input of the Document Transformer is the outputs of the export model and the output of the Document Transformer is the header summary. Another transformer approach [19] takes as input conversations of an encounter, i.e. human dialogues; it synapses them and produces abstract summaries. There is also a comparison of three pre-trained models based on transformers [20]. The input is news articles from the web and the output is an abstract summary of them. The pre-trained models used in this project are BART, PEGASUS and T5. After fine-tuning the models they seem to give satisfactory results and fluent summaries. The evaluation of each model is done with ROUGE, concluding that the T5 model surpassed all other models. Synapses of articles based on transformers are produced either by taking the dataset from the Wikihow knowledge base [21], or from web documents by leveraging relevant information from social media [22]. The pre-trained BERT and T5 models are used to produce the summaries.

In the field of social media, research is now oriented towards generating summaries of their content based on transformers. Research summarizing user comments under each post is in a particularly early stage. The content of both posts and user comments on social media presents peculiarities that only transformer models show they can handle. Most research

works have as input sources the platforms Twitter and Reddit [20, 23, 25], as well as Chinese social media platforms like "Sina Weibo" [24]. The summaries that are created concern the events, that is, the posts and less the comments of the users. They are mainly based on the pre-trained BERT model on the encoder side and a non-pre-trained transformer [23, 25], or other technologies [24] on the decoder side. The pre-trained T5 model, while applicable to many natural language tasks, finds limited application in generating summaries related to user comments [20].

At the same period an abstract summarization model was presented [26] which used a transformer model to generate individual sentence summaries of review texts. A combination of the Universal Sentence Encoder, statistical methods, and a graph reduction algorithm was then used to select the most relevant sentences to best represent the entire text in the summary.

Recently, Prompt Engineering has emerged as an innovative and powerful technique for improving the performance and adaptability of language models. This technique has its roots in LLMs who due to their ability to learn language patterns and structures have proven invaluable in various NLP tasks. However, due to the significant time and computational resources required to train these models, researchers turned to prompt engineering. New frameworks are being introduced, such as OpenPrompt [27], which conducts prompt-learning over PLM. It supports a combination of tasks (sort and generate), PLM (MLM, LM and Seq2Seq) and prompt modules. There are also introduced mechanisms for generating and improving summaries that use an entity chain intermediate scheme [28]. The summary is generated as a function of the entity chain and the input. Prefix-tuning is a novel idea to deal with the large size of today's LM changes [29]. Unlike fine-tuning which changes all model parameters and therefore needs to save the full copy for each task, prefix tuning freezes the language model parameters, modifying only a small task-specific continuous vector (called the prefix). Based on the prefix, personalization of needs can be achieved. Thus, a separate prefix can be created for each user according to their data and needs, thus allowing the production of specialized text. Thus, a separate prefix can be created for each user depending on their data and needs, thus enabling specialized text production. A variation of this idea is "prompt-tuning" [30] instead of "prefix-tuning". Finding that large models are expensive to share and maintain, reusing a frozen model for multiple downstream tasks can reduce this overhead.

Prompt engineering is experiencing great success in the healthcare field as traditional machine learning and deep learning methods did not solve NLP tasks in the medical field very well [31]. Prompt-based learning is a new promising area in the NLP field that aims to fill the gaps of the complexity of existing techniques and meet the complexity needs of natural language understanding and processing [32].

This paper is a presentation of a system that generates summaries of user comments on social media posts using

transformer models. The rest of the paper is organized into the following sections. Section 2 presents text summarization methods as well as the importance of summarizing social media content. Section 3 explains the importance of transformers against previous deep learning methods and gives a basic description of transformer model architecture. There is also a presentation of the methodology for the development of a model based on the pre-trained transformer models. The results obtained so far are shown in section 4. Consequently, Section 5 provides a discussion of the results and observations on model training, as well as future research perspectives. Finally, Section 6 presents the conclusions of the present research work.

2. Text Summarization Methods

The aim of text summarization is to convert a large text document into a shorter one by preserving the critical information and ensuring the meaning of the text. Due to large amount of data available, it is almost impossible for anyone to read all the comments generated by the users under a post. It is therefore necessary to properly code and program machines so that they can create coherent summaries just like humans. The process of text summarization done by machines or artificial intelligence programs is known as "Automatic Text Summarization". Automatic text summarization is the task of generating a concise and orderly summary while preserving the key information content and overall meaning [33].

There are three main text summarization approaches: extractive, abstractive, or hybrid [34]. Each approach is applied using different methods. This section will provide a detailed overview about each of these approaches along with the methods of each approach in the literature.

2.1. Extractive Summarization

In the extractive summarization the summarizer extracts important words and phrases from the original text and, gathering them together generates the summary. The words are extracted as they are in the original document with a slight rearrangement to give a structured sentence [Figure 1](#). According to El-Kassas et al. [34] the extractive summarization process consists of the following steps:

Creating a suitable representation of the input text for the text analysis purpose.

Scoring of sentences: based on the input text representation, statistical methods are used to score the frequency of occurrence of words or phrases.

Extraction of high-scored sentences: after finding the sentences or words with the highest score, a selection of the most important sentences is made to create the summary. An important issue that arises here is deciding the length of the abstract. There are length trimming or thresholding mechanisms to limit the size of the abstract and keep the same order of the generated sentences as the input text [35].

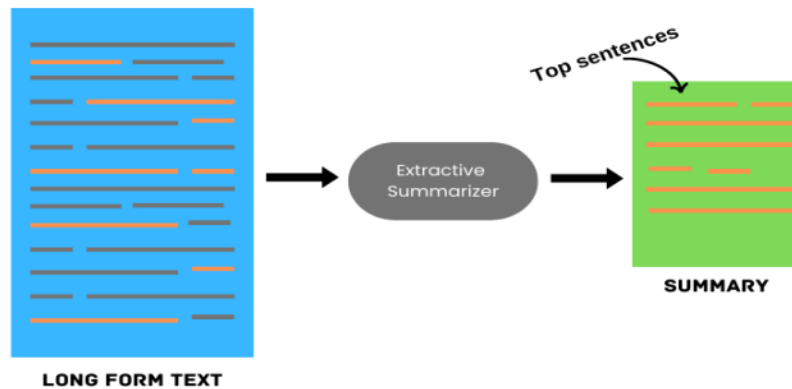


Figure 1. Extractive Text Summarization.

Advantages: The extractive approach is a simple and quick process. Furthermore, due to the use of exact sentences or words from the original text, it is noted for its accuracy. The sentences of the generated summary use the terminology of the original text, giving the reader accurate information.

Disadvantages: The extracting approach does not adopt the way humans make summaries, so it has several disadvantages:

- a) Redundancy of data without checking whether the generated sentences repeat the same information elsewhere in the generated summary text.
- b) Long sentences that lose control of the meaning and deviate from what the person is used to reading.
- c) Lack of semantics and coherence, since the produced text extracts entire sentences and words from the original text and places them in the final one. The selected sentences may be grammatically and syntactically correct but may not clearly convey the meaning of the original text.
- d) Dissemination of information. Information is scattered within sentences leading to the loss of important information or conflicting concepts. This problem is pronounced in short abstracts and tends to diminish when the abstract is long enough.

2.2. Abstractive Summarization

In contrast to extractive summarization, summaries generated by abstractive summarization are more human-like. This kind of summarization extracts the meaning of the original text and, using new words and phrases, creates a new shorter text that looks completely different but retains the meaning of the original [36]. This approach is more complex and sophisticated because it does not copy sentences or phrases from the original text, but uses NLP methods to understand the main concepts of the input text and relies on them to generate new sentences that make up the final output text *Figure 2*. The abstractive summarization process includes the two following tasks: a) building an internal semantic representation and b) generating a summary using natural language generation techniques to create a summary that is closer to the human-generated summaries [34]. Researchers involved in creating abstractive summaries following either linguistic or semantic approaches. Recently research on abstractive summarization has been inspired by deep learning and neural networks. More precisely Transformers and encoder-decoder models are being believed as smoother and also convenient in adjusting parameters automatically.

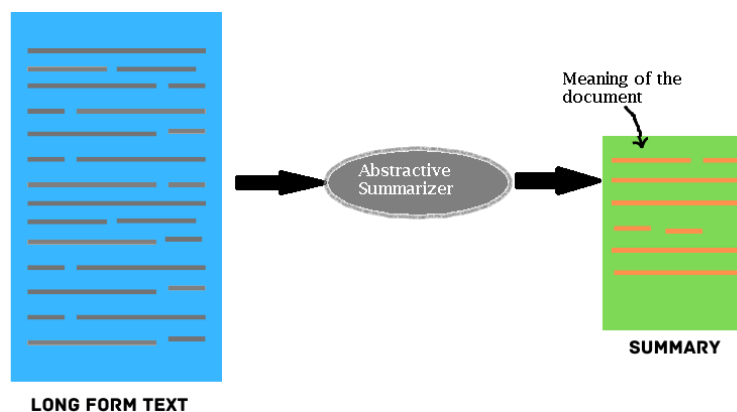


Figure 2. Abstractive Text Summarization.

Advantages: Summaries generated with this technique are better than the extractive ones and more human like. That is because they use different words/phrases that do not belong to the original text, based on semantics and the use of paraphrasing, compression, or fusion. Abstractive summaries overcome the problem of repetition. Compared to extractive, abstractive summaries can further reduce the length of the generated text while also avoiding redundancies.

Disadvantages: Abstractive summarization is difficult to implement since it is required a deep understanding of both the language and the text. Producing an informative, fluent, and readable summary remains a difficult task [37]. To achieve this task natural language generation techniques are needed which is still a growing field. The abstractive summarization faces difficulties and points that need to be improved. First of all, a detailed representation and understanding of the original text is needed to retain its meaning

and then render it in new words. Many times in the final text the same words are repeated in several places. In other words, there is no verbal richness. In addition, non-vocabulary words are difficult to deal with. In general the field of abstractive summarization is a growing field because the automatic summarizers do not have the verbal flexibility that the human mind presents.

2.3. Hybrid Summarization

It is a combination of extractive and abstractive summarization, Figure 3. The processing procedure consists of two stages: a) sentence extraction, thus the more important sentences of the input text are extracted such as in extractive summarization process procedure and b) the summary is generated following the abstractive summarization techniques.

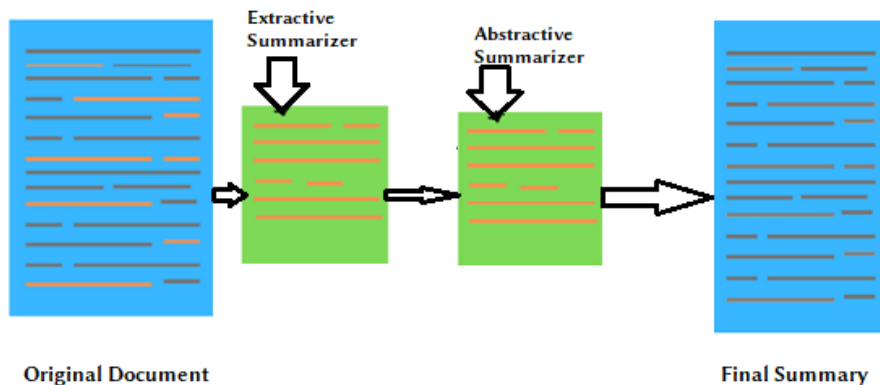


Figure 3. Hybrid Text Summarization.

Advantages: It combines the advantages of both extractive and abstractive summarization approaches.

Disadvantages: The generated summary is not as qualitative as the purely abstractive summary because it is not based on the original text but on the text extracted from the key sentences.

2.4. Social Media Summarization Significance

Social media's role of in public life is important as users' posts and comments shape public opinion on different but important issues concerning politics, economy and society. Social Media data excel in extracting sentiments and patterns of social behavior that can be used for social research, business decisions or government policy making. People tend to express themselves more openly in a relatively safe conversation environment with strangers. The expression of political views, psychological and social support for various causes by social assistance groups, is expressed more openly on the Internet than in the offline world. This is due to the anonymity provided by the internet, the implicit trust in the privacy of

communication and the absence of racial stereotypes. For these reasons, the data that circulates on social media is more representative of people's real opinions than in most offline interactions.

Social behaviors and emotions can be recorded and analyzed with statistical and graphical representations. Such representations can effectively capture the information related to various specific parameters from big social media data. But a text summary aims to capture the information related to the contents of various topics and present a coherent overview of those topics. For example, a statistical representation might rate a movie as good or a product as reliable on a rating scale of 0-5 [36]. But a text synopsis can actually give an overview of a movie's theme, a product's performance, or its technical problems, thus making a more specific review.

The sheer volume of information exchanged on the Internet and social media makes it imperative to create summaries to keep readers informed in an accurate and timely manner. Therefore creating summaries seems crucial because when an event breaks out, a huge number of posts and comments flood

the social network and most of them contain redundant and repetitive information, resulting in confusion for readers.

The area of social media summarization has been the focus of research recently. Social media's content remains challenging because of its specificity. Unlike official documents, social media content presents the following challenges:

- 1) It is informal, and ill-formed in terms of the grammar, structure and formality of Natural Language.
- 2) It is full of abbreviations, special characters, emoticons and slang expressions.
- 3) It lacks lexical richness, due to its brevity. Most textual content on social media is in the form of tweets, short comments and footnotes often represented by images/videos and lacking in linguistic richness. The present work trying to manage these peculiarities focuses on abstractive transformer-based summarization techniques.

3. Transformer Architecture and Methodology

3.1. Transformer Model

Transformer models were first proposed in the paper "Attention Is All You Need" [5]. Based on transfer learning, these models are increasingly attracting the interest of researchers as they can handle text-sequential contexts extremely well. So far they have modelled a state-of-the-art architecture in various NLP tasks such as translation, text generation, summarization, query answering, classification and sentiment analysis. The main reasons that research has focused on transformer models are the following three:

- a) The results they have given on many natural language tasks using sequential data are excellent.
- b) Unlike recurrent neural networks (RNNs) and convolutional networks that process words sequentially (one after the other), transformers support parallel processing. By replacing repetition with attention. Therefore, the calculations can be done simultaneously. This is very important because it ensures faster processing and saves computing resources.
- c) The architecture allows them to better handle long-range dependencies, which leads to better quality models. By having the attention mechanism they achieve the solution to the problem of RNN networks with long memory (LSTM) and CNN, which is mainly the inability to model larger sequences without data loss.

The transformer model architecture is an outgrowth of the encoder-decoder architecture in RNNs for handling sequence-to-sequence (seq2seq) tasks with an attention mechanism. The essential change brought by the transformers is that they eliminated the sequence factor, which allows greater parallelization and reduces training time.

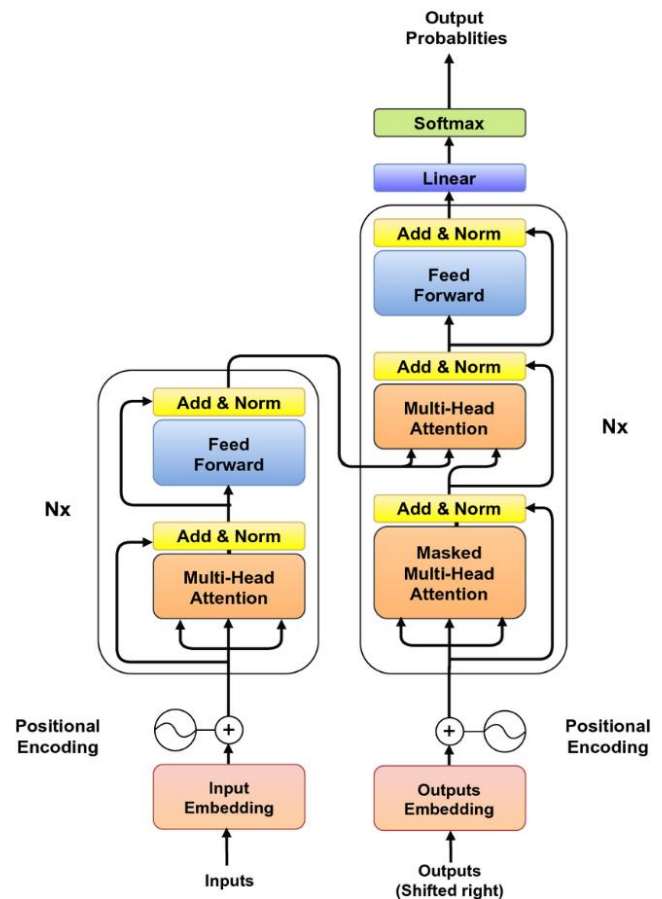


Figure 4. The Transformer Architecture.

The traditional transformer model on which the design of the present system is based operates with a encoder-decoder architecture that is widely used in the development of all neural networks. These architectures can handle variable-length sequences in an excellent way. For this reason, they are considered best suited for problems related to natural language understanding and text generation, such as translation and summarization. The traditional transformer model consists of two major building blocks such as Figure 4: an encoder (on the left) and a decoder (on the right). The model consists of two building blocks: the encoder and the decoder. Perpendicular to the two building blocks consists of a stack of N_x identical layers (in the original paper $N_x = 6$). Since the model does not contain any iteration or convolution, it has an extra layer of positional coding at the bottom of the encoder and decoder stacks to exploit the order of the sequence. Each layer mainly consists of Multi-Head Attention and feed forward layers. A variable-length sequence is used as input to the encoder, which is then converted into a numerical representation by extracting the important information from the input. This numeric representation is of fixed length. The decoder in turn maps the fixed-length encoded representation and converts it into a variable-length sequence to produce the output.

3.2. Pre-trained Transformer Models and Pipelines

For the construction of the system presented here, it was decided to use pre-trained models rather than building a transformer model from scratch for the following reasons: a) pre-trained models, if carefully and comprehensively pre-processed the data, give much better results, b) by providing an extensive learning base pre-trained models can tune many different datasets and c) it is easy to create new models with a small change in training and fine-tuning, leading to faster results [20].

A wide variety of pre-trained transformer models (PTLMs) are available today achieving excellent results in all natural language tasks. PTLMs vary in architectures and pre-training tasks. There are many models whose architectural backbone is Transformers, but some train only the encoder, such as BERT and UniLM, while others train only the decoder, such as GPT. Therefore, the knowledge of the special characteristics of each PTLM is important in order to make a correct choice of the appropriate one. In addition, the following criteria have been set:

- The model should accept text as input and produce text as output.
- The selected model should be able to produce abstractive summaries and should be based on the Encoder-Decoder architecture [38].
- All the pre-trained models can fine-tune the target task according to the constraints of the existing datasets.

The HuggingFace hub [39] is an open source library that provides a huge number of pre-trained models as well as datasets for a wide range of different NLP tasks. These models can be used to predict a summary and then accurately fine-tune any data set. As already mentioned above, because choosing the appropriate model is a difficult process, a first approach may be the use of pipelines. Pipelines are the fastest, easiest and most efficient way to use different pre-trained models, which can be applied to many NLP tasks. For the purposes of this work, the pipelines that meet the above criteria are three: BART [40], T5 [41] and PEGASUS [43]. Thus, a comparison of three pre-trained transformer models (T5, BART and PEGASUS) that met the above criteria was first performed to generate summaries of the data set used for this work [44], Figure 5.

Model T5 showed the best behaviour of the three. T5 stands for "Text-to-Text Transfer Transformer" [41] and is based on encoder-decoder architecture. It converts all NLP problems as text-to-text problems and can be trained or fine-tuned on either supervised or unsupervised data. T5 is suitable for any NLP task such as Translation, Language Induction, Information Extraction and Summarization, and so it is characterised as task agnostic. Its training procedure is based on teacher forcing. It has an input sequence and a target sequence. Additionally, T5 is an early LLM that has one of the best performances in natural language tasks [42].

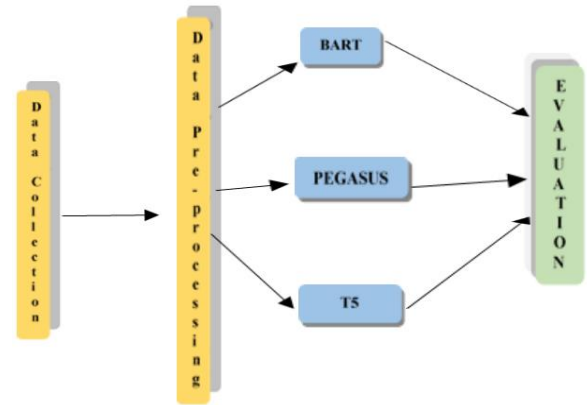


Figure 5. Pre-trained model evaluation.

3.3. Methodology

The particularities of data content in social media have already been mentioned, making the creation of summaries particularly tempting. These are not formal texts, articles or documents rich in grammar and linguistics. On the contrary, these are short texts consisting of abbreviations, slang expressions, special characters and emoticons. In addition, the redundant and repetitive information they contain lead readers to confusion. Thus, to improve the performance of the summary generation model, the greatest weight is given to pre-processing.

There are many dataset availability platforms for natural language processing but specific social media datasets suitable for summarization have not been widely available. Finding and downloading the appropriate datasets has been quite a complicated process since social media platforms limit data downloads. This project uses a dataset of Facebook news posts accompanied by user comments below each post. The necessary transformations were applied to the data to remove unnecessary elements and to preserve and group the useful ones. The raw data has 7 columns namely "created_time", "from_id", "from_name", "message", "post_name", "post_title", "post_num" and 1,781,576 rows. The adjustments made were to retain 3 of the 7 columns. Thus, the existence of 'post_title' and 'post_number' which are identifying elements of a specific post and help to group posts and events was deemed necessary. In addition, the 'message' column was used as the basic information because it contains the user comments to be summarized. The approach presented in this paper consists of the following steps Figure 6:

- 1) Data collection
- 2) Pre-processing.
- 3) Topic-based data grouping
- 4) Feed the encoder with the text lists
- 5) Pass data to the decoder
- 6) Generate summaries
- 7) Summary validation

After the data set has been taken and transformed in a way suitable for generating the summary, the next very important

step in properly fitting the model is preprocessing. In the sense of pre-processing all procedures that lead to the cleaning of the data from unnecessary and useless elements that would lead the model to issue unwanted results. This stage was implemented with the help of NLTK and regex python packages. So, punctuation, special characters, emoji strings, and dashes were removed, as well as NULL values. The data was then grouped by discussion topic, as the goal of the system is to generate a summary of user comments related to a particular post topic. So based on the title of each post, the data dictionary is reconfigured to isolate user comments for each post into separate lists of different sizes. The input sequence is formatted and processed to convert each word into a unique numeric identifier using `body_input_ids` with

`body_attention_masks`. At the embedding layer, the transformer transforms the input tokens into vectors using learned embeddings. There are two main parts that make up the encoder: a multi-head attention mechanism followed by normalization and a feed-forward neural network. The Multi-head Attention mechanism is based on a dot attention scale that generates a vector for each input word. Having 8 attention heads it has the ability to generate multiple vectors for the same word. This is very important because it helps the model to capture different representations of word relationships in the sentence creating different attention matrices. These matrices are finally summed and passed through a linear layer to create a single reduced matrix.

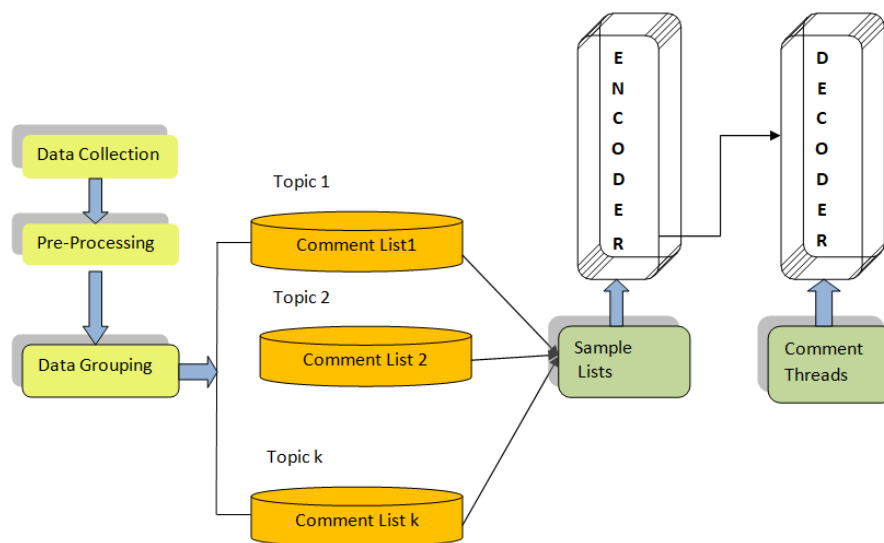


Figure 6. System's methodology.

Two inputs are required on the decoder side: the encoder output (`summary_input_ids`) and the right-shifted output text (`generated_ids`). Then the multi-head attention mechanism is applied twice with one of them being "masks" (`summary_attention_masks`). The dictionary in the last level of the decoder should be the same size as the target dictionary. Finally, by applying the softmax function, the probability of each word being present in the output is indicated.

4. Results and Validation

As already mentioned, in this work a dataset of posts and user comments from Facebook was used. It was developed by applying the traditional transformer model to a computer with an NVIDIA GeForce 4070 GPU with 12 BG of RAM. Due to the large amount of data (1,781,576 rows) and the available computing resources for brevity used 1/3 of the obtained data set, after it was pre-trained. The new data, grouped by topic, was used for training. The dataset was divided into three

subsets: Train 80%, Validation 20% and Test (validation/2=) 10%. The current model is trained on the pre-trained 12-level T5-base with 12 attention heads and network feed depth as 3072. The AdamW optimizer was used with 1e-3 learning rate. The dropout rate was set to 0.1. The model needed 12 epochs to be trained with batch size set to 64.

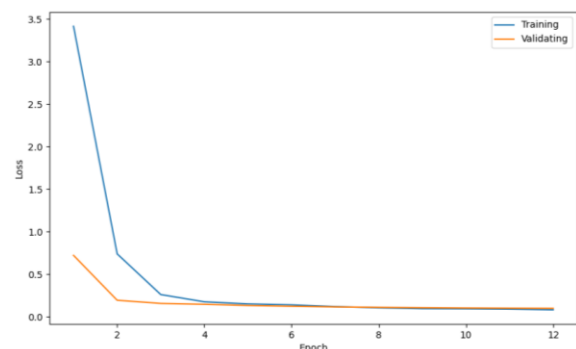


Figure 7. Train and Validation Loss.

To evaluate the performance of the model, the loss was calculated. Loss is considered one of the most important factors in deep learning models as it measures the error produced by the model. Initially the dataset was divided into three subsets: Train, Validation and Test. The training and validation loss was then measured for all samples. As is evident from [Figure 7](#), a gradual decrease of both metrics is observed with the gap between validation and train loss shrinking after each epoch. This leads to the conclusion that the model learns from the data and can produce satisfactory results because it shrinks the regularization loss (model weights), leading to a small difference between validation and train loss.

Table 1. Rouge Score Results.

	Rouge 1	Rouge 2	Rouge L
P	0.520	0.514	0.519
R	0.854	0.849	0.854
F	0.646	0.640	0.645

The Recall-Oriented Understudy for Gisting Evaluation, ROUGE [45] is the most popular metric so far, for evaluating automatic summarization quality. It measures the n-gram overlap between the generated summary and the reference summary. At this stage, the comparison Rouge makes is between the system-generated summary and each comment in the thread. The results are shown in [Table 1](#). Specifically, the precision, recall and F1 score are calculated. Accuracy gives a ratio of words suggested by the predicted abstract to those actually appearing in the reference abstract (0.514-0.520). Recall gives a ratio of words in the reference abstract captured by the predicted abstract (0.849-0.854). Finally f1 represents the harmonic mean score of the two previous ratios (0.640-0.645).

5. Discussion

Rouge ratings range from 0-1 with 1 having the best price lead. Therefore, the closer the measurements are to one, the better the summary produced. The higher the score in Rouge means the system captures the most important information to include in the summary. This does not necessarily mean that the generated summary is of high quality, as it may contain biased text. The evaluation of the quality of the generated text is considered a complex process where the dimensions and limitations of each system should be taken into account.

At this stage of research, to evaluate the generated summaries ROUGE scored by comparing the generated summary to each comment in the thread. But the comments posted on social media are informal and lack lexical richness. A quality

abstract should have lexical richness and coherence. A better approach would be to compare the system-generated summary with the human-generated summary. ROUGE only works on n-grams. A score of 1 represents the perfect summary. But this would only happen if both summaries had the same n-grams. Additionally, ROUGE performs better on models that produce an extractive summary. Since the presented model focuses on paraphrasing, thus creating abstractive summaries, the research carried out will be extended to optimize the results with more modern techniques and tools. In this research, datasets from Facebook corresponding to specific posts and user comments below them were used as input data.

6. Conclusion

The digitization of our daily lives through smart technology has made social media the dominant way of informing, expressing social, psychological, economic and political beliefs. It is the most direct way of exchanging views and shaping public opinion. The amount of information posted daily on social media is growing exponentially, making it practically impossible to read. The summary of this information is an extremely important tool for correct, timely and accurate information to readers. The number of comments is proportional to the importance of the post and is often so large that it is difficult and time-consuming to read all of them. User comments are important to read, both by individual readers and by the users or groups of users who made the post. They express public opinion that both those who make the post and those who read it are important to know. Social media content is an area of challenge: a) it offers different types of information, b) online conversations and conversation threads are informal, experiencing language deficiency and lack of lexical richness. c) Comments are full of abbreviations, slang expressions, special symbols, hashtags and emoticons. For all the above reasons, creating summaries from user comments is a complicated task. In this article, various social media summaries development technologies have been reported with the transformer models dominating the recent research. However, each technology is implemented under different assumptions and limitations.

Transformer-based models have a talent for natural language processing and have given excellent results in creating summaries. However, the field of social media remains unexplored due to the nature of its content. The current approach focuses on developing a system that generates an abstract summary of user comments under a social media post. Specifically, transformer-based encoder-decoder architecture is applied using a dataset of user comments on Facebook posts as input. Three different pre-trained models -BART, T5, PEGASUS- were compared using ROUGE metrics, concluding that T5 gives the highest performance on the dataset used. In addition, the T5 is one of the first models in the LLM category and has given excellent results in language production.

Measuring the performance of the system led to two main conclusions: a) the learning rate of the model seems very good according to the given data set. Initially due to lack of learning the values of training and validation sets start with high values and their difference is large. As they are trained, a reduction in errors and a smoothing of the curves is observed. This leads to the conclusion that the model is trained correctly and can give correct results. b) Rouge metrics, the generated summaries are quite satisfactory. As with all research work, however, there are areas for improvement. ROUGE performs better on models that produce an extractive summary. Since the presented model focuses on paraphrasing, thus creating abstractive summaries, the research carried out will be extended to optimize the results with more modern techniques and tools. Additionally, instead of user threads as reference summary, a human generated summary should be used leading to human texts. Finally, the system should be applicable to all social network platforms.

Abbreviations

NLP	Natural Language Processing
RNN	Recurrent Neural Network
CNN	Convolutional Neural Network
PR	Phrase Reinforcement
TF-IDF	Term Frequency-Inverse Document Frequency
DTM	Decomposition Topic Model
GDTM	Gaussian Decomposition Topic Model
LSTM	Long Short-Term Memory
BERT	Bidirectional Encoder Representations from Transformers
BART	Bidirectional and Auto-Regressive Transformer
T5	Text-to-Text Transfer Transformer
LCSTS	Large-scale Chinese Short Text Summarization
LLM	Large Language Model
PLM	Prompt Learning Model
MLM	Masked Language Model
LM	Language Model
Seq2Seq	Sequence to Sequence
PTLM	Pre-Trained Language Model
UniLM	Unified Language Model
GPT	Generative Pre-trained Transformer
ROUGE	Recall-Oriented Understudy for Gisting Evaluation

Acknowledgments

The publication fees were totally covered by ELKE at the University of West Attica.

Author Contributions

Afrodite Papagiannopoulou: Investigation, Methodology, Validation, Writing – original draft

Chrissanthi Angeli: Conceptualization, Project administration, Supervision

Data Availability Statement

The data supporting the outcome of this research work has been reported in this manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Gupta, S. and Gupta, S. K. Abstractive summarization: An overview of the state of the art. *Expert Systems with Applications* 121, 2019, pp. 49–65. <https://doi.org/10.1016/j.eswa.2018.12.011>
- [2] Luhn, H. P. The Automatic Creation of Literature Abstracts. *IBM Journal of Research and Development*, vol. 2, no. 2, Apr. 1958, pp. 159-165, <https://doi.org/10.1147/rd.22.0159>
- [3] Suleiman, D., A. Awajan, A. Deep Learning Based Abstractive Text Summarization: Approaches, Datasets, Evaluation Measures, and Challenges. *Mathematical Problems in Engineering*, 2020, <https://doi.org/10.1155/2020/9365340>
- [4] Gupta, V., Lehal, G. S. A Survey of Text Summarization Extractive techniques. *Journal of Emerging Technologies in Web Intelligence*, 2010, pp. 258–268, <https://doi.org/10.4304/jetwi.2.3.258-268>
- [5] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. Attention Is All You Need In 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA., June 2017. <https://arxiv.org/abs/1706.03762>
- [6] Sharifi, B., Hutton, M-A. and Kalita, J. (2010) “Summarizing Microblogs Automatically”. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 685–688, Los Angeles, California. Association for Computational Linguistics.
- [7] Sharifi, B., Inouye, D., and Kalita, J.K. (2014) “Summarization of Twitter Microblogs”. *The Computer Journal*, Volume 57, Issue 3, March 2014, Pages 378–402, <https://doi.org/10.1093/comjnl/bxt109>
- [8] F. Amato, F. Moscato, V. Moscato, A. Picariello, G. Sperli’, “Summarizing social media content for multimedia stories creation”. *The 27th Italian Symposium on Advanced Database Systems (SEB 2019)*. <https://ceur-ws.org/Vol-2400/paper-40.pdf>
- [9] F. Amato, A. Castiglione, F. Mercorio, M. Mezzanzanica, V. Moscato, A. Picariello, G. Sperli’ “Multimedia story creation on social networks,” *Future Generation Computer Systems*, 86, 412–420, 2018, <https://doi.org/10.1016/j.future.2018.04.006>

- [10] J. Bian, Y. Yang, H. Zhang, T. S. Chua, "Multimedia summarization for social events in microblog stream," *IEEE Transactions on Multimedia*, 17(2), 216–228, 2015, <https://doi.org/10.1109/TMM.2014.2384912>
- [11] D. Chakrabarti, K. Punera, "Event Summarization Using Tweets". *Proceedings of the International AAAI Conference on Web and Social Media*, 5(1), 66–73. <https://doi.org/10.1609/icwsm.v5i1.14138.2011>
- [12] Chong, F., Chua, T., Asur, S. (2021) "Automatic Summarization of Events from Social Media". *Proceedings of the International AAAI Conference on Web and Social Media*, 7(1), 81–90. <https://doi.org/10.1609/icwsm.v7i1.14394>
- [13] Gao, S., Chen, X., Li, P., Ren, Z., Bing, L., Zhao, D. and Yan, R. (2019) "Abstractive Text Summarization by Incorporating Reader Comments". In *The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*, 33(01): 6399–6406. <https://doi.org/10.1609/aaai.v33i01.33016399>
- [14] Liang, Z., Du, J. and Li, C. (2020) "Abstractive social media text summarization using selective reinforced Seq2Seq attention model," *Neurocomputing*, 410, 432–440, <https://doi.org/10.1016/j.neucom.2020.04.137>
- [15] Wang, Q. and Ren, J. (2021) "Summary-aware attention for social media short text abstractive summarization," *Neurocomputing*, 425, 290–299, <https://doi.org/10.1016/j.neucom.2020.04.136>
- [16] Bhandarkar, P., Thomas, K. T. (2023) "Text Summarization Using Combination of Sequence-To-Sequence Model with Attention Approach", Springer Science and Business Media Deutschland GmbH: 283–293, 2023, https://doi.org/10.1007/978-981-19-3035-5_22
- [17] Gupta, A., Chugh, D. and Katarya, R. (2022) "Automated News Summarization Using Transformers", In *Sustainable Advanced Computing*, 2022, Volume 840. ISBN: 978-981-16-9011-2. <https://arxiv.org/pdf/2108.01064>
- [18] M. H. Su, C. H. Wu, H. T. Cheng, "A Two-Stage Transformer-Based Approach for Variable-Length Abstractive Summarization," *IEEE/ACM Transactions on Audio Speech and Language Processing*, 28, 2061–2072, 2020, <https://doi.org/10.1109/TASLP.2020.3006731>
- [19] D. Singhal, K. Khatter, A. Tejaswini, R. Jayashree, "Abstractive Summarization of Meeting Conversations," in *2020 IEEE International Conference for Innovation in Technology, INOCON 2020*, Institute of Electrical and Electronics Engineers Inc., 2020, <https://doi.org/10.1109/INOCON50539.2020.9298305>
- [20] Ivan S. Blekanov, Nikita Tarasov and Svetlana S. Bodrunova. 2022. Transformer-Based Abstractive Summarization for Reddit and Twitter: Single Posts vs. Comment Pools in Three Languages. *Future Internet* 14, 69. <https://doi.org/10.3390/fi14030069>
- [21] A. Pal, L. Fan, V. Igodifo, Text Summarization using BERT and T5. https://anjali001.github.io/Project_Report.pdf
- [22] M. T. Nguyen, V. C. Nguyen, H. T. Vu, V. H. Nguyen, "Transformer-based Summarization by Exploiting Social Information," in *Proceedings - 2020 12th International Conference on Knowledge and Systems Engineering, KSE 2020*, Institute of Electrical and Electronics Engineers Inc.: 25–30, 2020, <https://doi.org/10.1109/KSE50997.2020.9287388>
- [23] Li, Q. and Zhang, Q. (2020) "Abstractive Event Summarization on Twitter". In *The Web Conference 2020 - Companion of the World Wide Web Conference, WWW 2020*, Association for Computing Machinery: 22–23, <https://doi.org/10.1145/3366424.3382678>
- [24] Z. Kerui, H. Haichao, L. Yuxia, "Automatic text summarization on social media," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, 2020, <https://doi.org/10.1145/3440084.3441182>
- [25] Tampe, I., Mendoza, M. and Milios, E. (2021) "Neural Abstractive Unsupervised Summarization of Online News Discussions". In: Arai, K. (Eds) *Intelligent Systems and Applications. IntelliSys 2021. Lecture Notes in Networks and Systems*, vol 295. Springer, Cham. <https://arxiv.org/abs/2106.03953>
- [26] Rawat, R., Rawat, P., Elahi V. and Elahi, A. (2021) "Abstractive Summarization on Dynamically Changing Text," *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2021, pp. 1158–1163, <https://doi.org/10.1109/ICCMC51019.2021.9418438>
- [27] Ding N, Hu S, Zhao W, Chen Y, Liu Z, Zheng H, et al. Open-Prompt: An Open-source Framework for Prompt-learning. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*; 2022. p. 105–13. <https://doi.org/10.18653/v1/2022.acl-demo.10>
- [28] Shashi Narayan, Yao Zhao, Joshua Maynez, Gonalo Simões, Vitaly Nikolaev, and Ryan McDonald. 2021. Planning with Learned Entity Prompts for Abstractive Summarization. *Transactions of the Association for Computational Linguistics*, 9: 1475–1492. https://doi.org/10.1162/tacl_a_00438
- [29] Xiang Lisa Li and Percy Liang. 2021. Prefix-tuning: Optimizing continuous prompts for generation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 4582–4597, Online. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.acl-long.353>
- [30] Brian Lester, Rami Al-Rfou, and Noah Constant. 2021. The Power of Scale for Parameter-Efficient Prompt Tuning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3045–3059, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.emnlp-main.243>
- [31] Wang, Jiaqi, Enze Shi, Sigang Yu, Zihao Wu, Chong Ma, Haixing Dai, Qiushi Yang, Yanqing Kang, Jinru Wu, Huawei Hu, Chenxi Yue, Haiyang Zhang, Yi-Hsueh Liu, Xiang Li, Bao Ge, Dajiang Zhu, Yixuan Yuan, Dinggang Shen, Tianming Liu and Shu Zhang. "Prompt Engineering for Healthcare: Methodologies and Applications." <https://arxiv.org/html/2304.14670v2>.

- [32] Liu, Pengfei, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi and Graham Neubig. "Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing." *ACM Computing Surveys* 55(2021): 1-35. <https://doi.org/10.1145/3560815>
- [33] Chua, F., Asur, S. Automatic Summarization of Events from Social Media. In *Proceedings of the International AAAI Conference on Web and Social Media, 2023*, 7(1), pp. 81-90. <https://ojs.aaai.org/index.php/ICWSM/article/view/14394>
- [34] El-Kassas, W. S., Salama, C. R., Rafea, A. A. and Monhamed, H. K. Automatic Text Summarization: A Comprehensive Survey. *Expert Systems with Applications*. (2020). 165, 113679. <https://doi.org/10.1016/j.eswa.2020.113679>
- [35] Wang, S., Zhao, X., Li, B., Ge, B. & Tang, D. Integrating extractive and abstractive models for long text summarization. *IEEE International Congress on Big Data (Big Data Congress)*, 2017, pp. 305-312. <https://ieeexplore.ieee.org/document/8029339>
- [36] Varma, V., Kurisinkel, J., Radhakrishnan, P. Social Media Summarization, Cambria, E., Das, D., Bandyopadhyay, S., Feraco, A. (eds) *A Practical Guide to Sentiment Analysis. Socio-Affective Computing*, vol 5. Springer, Cham., 2017, pp 135–153 https://doi.org/10.1007/978-3-319-55394-8_7
- [37] Lin, H., & Ng, V. Abstractive Summarization: A Survey of the State of the Art. In *Proceedings of the AAAI Conference on Artificial Intelligence, 2019*, 33(01), pp. 9815-9822. <https://doi.org/10.1609/aaai.v33i01.33019815>
- [38] K. Pipalia, R. Bhadja, M. Shukla, "Comparative analysis of different transformer based architectures used in sentiment analysis," in *Proceedings of the 2020 9th International Conference on System Modeling and Advancement in Research Trends, SMART 2020*, Institute of Electrical and Electronics Engineers Inc.: 411–415, 2020, <https://doi.org/10.1109/SMART50582.2020.9337081>
- [39] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. Le Scao, S. Gugger, M. Drame, Q. Lhoest, A. M. Rush, "HuggingFace's Transformers: State-of-the-art Natural Language Processing," 2019. <https://doi.org/10.18653/v1/2020.emnlp-demos.6>
- [40] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V. and Zettlemoyer, L. (2019) "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension". In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics. <https://aclanthology.org/2020.acl-main.703>
- [41] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narag, S., Matena, M., Zhou, Y., Li, W. and Liu P. J. (2021) "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer". In *The Journal of Machine Learning Research*, Volume 21, Issue 1, 2019. ISSN: 1532-4435. <https://dl.acm.org/doi/abs/10.5555/3455716.3455856>
- [42] Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Barnes, N., & Mian, A. S. (2023). *A Comprehensive Overview of Large Language Models*. <http://arxiv.org/pdf/2307.06435>
- [43] Zhang, J., Zhao, Y., Saleh, M. and Liu, P. J. (2020) "PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization". In *ICML'20: Proceedings of the 37th International Conference on Machine Learning*, July 2020, Article No.: 1051, Pages 11328–11339. <https://dl.acm.org/doi/abs/10.5555/3524938.3525989>
- [44] Rawat A., and Singh Samant, S. (2022) "Comparative Analysis of Transformer based Models for Question Answering". 2nd International Conference on Innovative Sustainable Computational Technologies (CISCT), Dehradun, India, 2022, pp. 1-6, <https://doi.org/10.1109/CISCT55310.2022.10046525>
- [45] Lin, C.-Y. (2004) "ROUGE: A Package for Automatic Evaluation of Summaries". In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain. Association for Computational Linguistics. <https://arxiv.org/abs/1803.01937>

Biography



Afrodite Papagiannopoulou is a PhD student at the University of West Attica in Athens Greece, Department of Electrical and Electronics Engineering. She received her MSc in Intelligent Knowledge Based Systems from the University of Essex U.K. in 1996. She is working as head of IT in the education department of the Greek Ministry of Education. Her research interests include artificial intelligence techniques and neural networks for natural language generation, social media, abstractive summarization and transformers.



Chrissanthi Angeli is a Professor at the University of West Attica, Athens Greece, Department of Electrical and Electronics Engineering. She holds an MSc in Intelligent Systems from the University of Plymouth U.K and a Phd in Intelligent Fault detection techniques from the University of

Sussex U.K. Her current research interests include AI techniques for fault detection/prediction, modelling and simulation, neural networks for natural language generation.

Research Field

Afrodite Papagiannopoulou: Artificial Intelligence, Neural Networks, Natural Language Generation

Chrissanthi Angeli: Artificial Intelligence techniques, Neural Networks, Natural Language Generation