Text Summarization Adaptive Models for Semantic Relevance Information: A Survey

Ajit Kumar Rout [1], Abhisek Sethy[*1], M Ram Kumar [2],Md Fasi Ahamed [3], S Mohan[4]

Department of Information Technology
GMR Institute of Technology, Rajam, Andhra Pradesh, India

Abstract: Text Summarization has long been a prominent field of research in the Artificial Intelligence. Text summarising is the method to extract the essential ideas and meaning from a given text and trying to turn them into a summary. Automatic summarization has evolved into a crucial technique for quickly and efficiently finding important information in large amounts of text. It is really challenging to summarize large amount of data but by utilising various Natural Language approaches, it can be accomplished with pre-trained models, a review of adaptive models for summarization is presented in this study such as 'DA-PN + Cover + MLO' model, Auto-Encoder (AE), Variational Auto Encoder (VAE), Extreme Learning Machine Auto-Encoder (ELMAE) and Text Rank algorithm. The main objectives of this study are to auto-summarize the text given by user and to estimate idea importance and choose the lines that will be contained in the conclusion that are the most important.

***Keywords**:* NLP, Summary, Encoder, ELM Auto-Encoder, VAE, Decoder and Text Rank Algorithm.

1. Introduction

Because of the growth of internet social media and user-generated content, the quantity and amount of digital files available on the web have greatly increased. NLP applications like text summarization will unquestionably have a big impact on how we live our lives. Text summarization systems have grown in importance provided by massive increase in text information produced by social networks, sensors, and news websites, among others. These systems allow users to quickly and easily understand a text without having to scroll through countless pages, which could save those hours of searching and help them concentrate on their intended use. Who has the time to read whole news articles, documents, and books to assess if they are useful or not in the age of digital media and ever-increasing publication. Among the trickiest and most fascinating areas of NLP is automatic text summarization (ATS). It describes the method of creating a succinct and cohesive summary of writing from many text sources[12], such as books, news articles, blog posts, research papers, emails, and tweets. Demand for automatic summarization technology is increasing as massive amounts of textual data become available.

To determine the semantic similarity between sentences, we take into account a variety of models like 'DA-PN + Cover + MLO' model, Auto-Encoder (AE), Variational Auto Encoder (VAE), Extreme Learning Machine Auto-Encoder (ELMAE)[13], Encoder and Decoder and

Text Rank algorithm.According to our analysis of the literature, the new method provides certain benefits over the earlier ones. In the initial stage, it represents implicit semantic relations rather than using explicit semantic linkages provided by outside lexical resources like WordNet[13]. Several semantic-based strategies have advanced in the recent times. WordNet is among the most often used English thesauruses, it has been extensively utilized to enhance a variety of NLP approaches, such as automated summarising. Therefore, a much reliable method that will identify the existing semantic links between various textual units is required. Second, instead of employing feature extraction techniques or domain knowledge, it automatically learns high-level features from data using unsupervised feature learning. Our investigation of text based summarising strategies describes that they are constructed on a bag-of-words form, which necessitates sparse and increased input data and it is difficult to capture semantic linkages between textual units[14].

2. Literature Review

To enhance ATS, research Zhang et al. in [1] has proposed improvements towards the design of the attention-based bi-directional LSTM model ('Bi-LSTM + Attention') and the attention-based sequence model ('Seq2Seq + Attention'). For the purpose of increasing the possibility of producing more accurate text summaries, which also improves the correlation with the source text, Firstly a novel enhanced semantic network (ESN) model has been proposed. This model analyses the closeness of the semantic meanings of encoder and decoder to maximise the semantic relevance during training. Second, a novel "DA-PN" system makes a pointer network (PN)-based model that uses decoder attention (DA) has been presented to deal with the issue of unregistered words. Thirdly, having been suggested to combine the complex "DA-PN" system with a multi-attention-integrated coverage method.

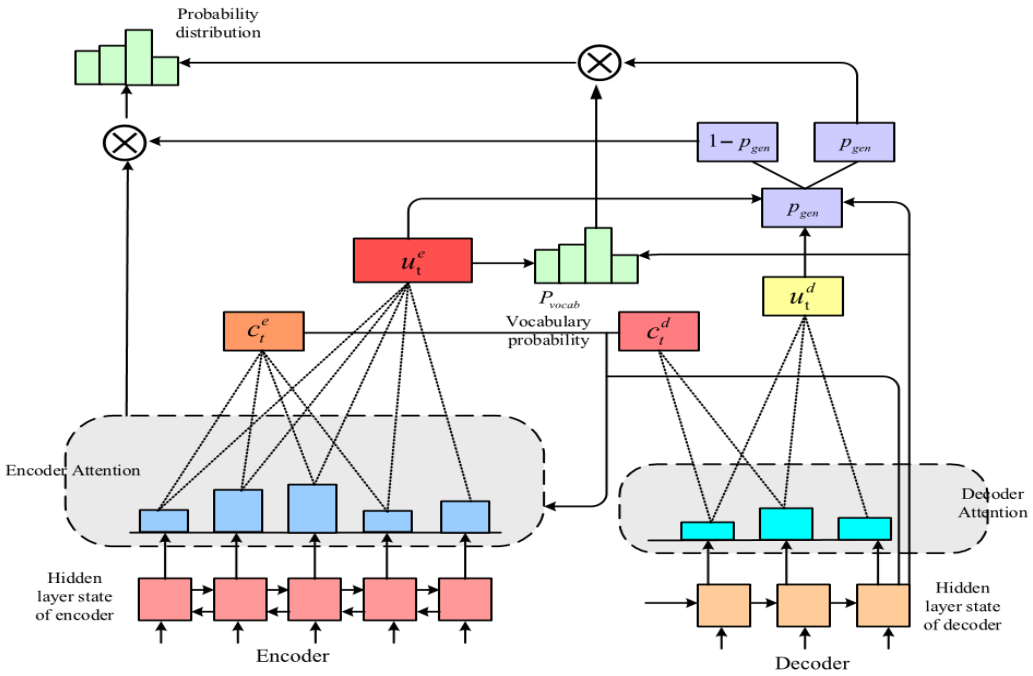

Figure 1. The Proposed 'DA-PN + Cover' model.

The attention as of the present time step is positively impacted in the resulting "DA-PN + Cover" model by using the attention distribution of the encoder and decoder previously, at every time step, that results in the identification of more accurate words for inclusion in the text summaries by avoiding repeated words depicted in Figure 1.

Table 1. Performance comparison of models

Model	ROUGE-1	Rouge-2	ROUGE-L
'Bi-LSTM + Attention' (baseline)	25.27	8.03	29.04
'Seq2Seq + Attention' (baseline)	27.07	9.03	30.56
ESN	**29.17**	**10.97**	**30.95**
'DA_PN'	**32.20**	**13.46**	**31.26**
'DA-PN + Cover'	**33.43**	**13.77**	**32.93**
'DA-PN + Cover + MLO'	**33.75**	**14.09**	**32.69**

Lastly, it has been suggested to incorporate a mixed learning objective (MLO) function into the "DA-PN + Cover" model to be able to stop the propagation of cumulative errors in generated text summaries. The 'DA-PN + Cover + MLO' model that was produced is the one that performs the best among the ATS models suggested in this study in Table 1. Due to the large number of unlabelled data sets available and not sufficient label data to train supervised models, unsupervised approaches to DL are preferable. Alami et al. in [2] had introduced a variety of unsupervised approaches to DL have been introduced to learn features from unlabelled data sets, the problem of a lack of data sets labelled is no anymore a problem. Auto-Encoder (AE), Variational Auto Encoder (VAE), and ELM-AE are a couple of illustrations of such models used in this study. Instead of using the Bag of Words representation in this paper, word2vec model utilised as the input for models. In addition, they unveiled a brand-new model based on ELM-AE for text summarization. Both the BOW and word2vec approaches used to training the model ELM-AE were studied for their effects. First of all, the output is always equal to the output in any auto encoder model. Finding the latent space is the problem at hand, though (compressed data). The decoder model is altered based on the error parameters. Mean square error is the error measure in use here. The Gaussian distribution with mean and SD as parameters must be utilised for the latent space in the variational auto encoder model. These parameters are used to produce the outcome in Figure 2.

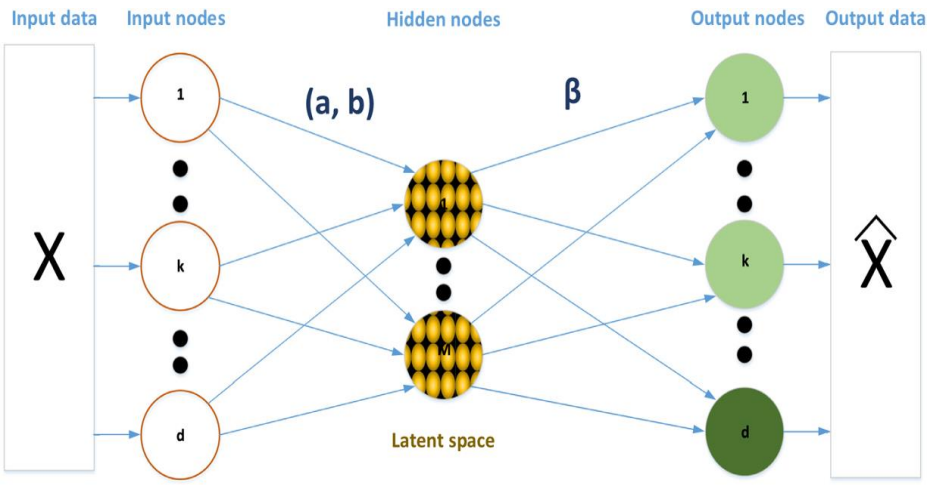

Figure 2. ELM-AE model

d > M: To compress latent space

d = M: Equal input and latent space dimensions

d < M: Enlarged representation

The three following conditions determine how the training settings for the ELM-AE are changed:

$$\beta = \left(\frac{I_M}{C} + H^T H\right)^{-1} H^T X$$

$$\beta = H^T \left(\frac{I_N}{C} + HH^T\right)^{-1} X$$

$$\beta = H^{-1} X$$

Finally, using a voting mechanism, we suggest three new ensemble techniques that aggregate the outcomes of many researched models.

In this study, Zhao in [3] introduced a model for creating short text summaries based on keyword templates is put forth in an effort to enhance how Chinese short texts are preprocessed for use in these activities.They used the classic encoder-decoder structure and tried to increase effectiveness of their suggested method in producing brief text summaries. The decoder uses eight Transformers stacked with random initialization, whereas the encoder uses BERT for initialization. Their model was given the name BSA. If the model divides sentences into categories using keywords, it is called BSA*. They demonstrated how to effectively construct brief text summaries using the pre-trained language model, and they demonstrated how to validate it using the abstractive method. This model can function as a springboard to raise the summary's quality and make the pre-trained language model more effective at producing concise text summaries. In order to train on a big corpus, the BERT model employs "masked language modelling" and "next sentence prediction" techniques. BERT is useful for an amount of downstream tasks in natural language processing since when used on them, it can be tweaked to give exceptional results in figure 3.

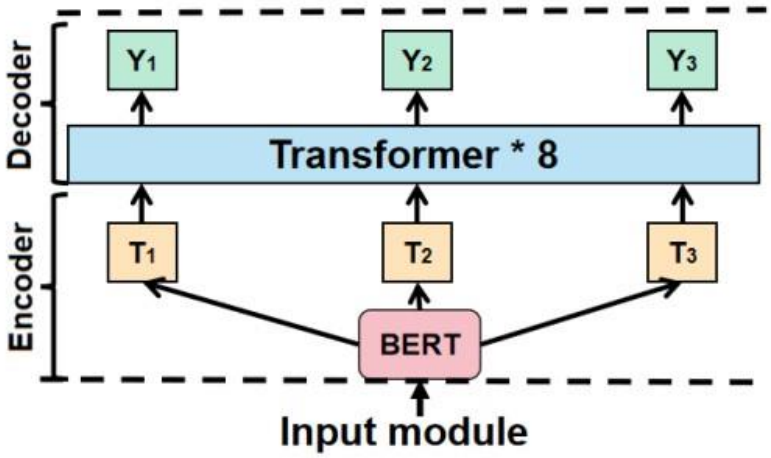

Figure 3.The BSA model's architecture structure. It includes a encoder (BERT) and an decoder (Transformer)

Zhao et al. [4] suggested model performs better with a 44.2 F-score for ROUGE-1, 28.9 ROUGE-2, and 39.2 for ROUGE-2. The proposed model is enhanced by 4.3, 7.4, and 1.3 compared to the RTCS model. Text summaries are meant to make the text simpler to read and comprehend, particularly for weak readers. Extractive and abstractive summary generation are the two types. From the original materials, key words or phrases are taken out by extractive summarization algorithms. It is more developed and simple when compared to abstractive approaches. The model uses the encoder to compress the original material into vectors as dense, and decoder uses the compressed vectors to create a summary. A top-notch summarising system must concentrate on the topic material of the content and the similarity of the original text and the summary. For text summarization, they suggest topic information fusion and semantic relevance based on fine-tuning BERT. As a contribution to the study, topic keywords are first extracted and combined with the source materials. Second, by directing the summary at the original document, the accuracy of the abstract is improved and determining the degree of semantic similarity among generated summary and the source content. To maximise the similarity score through training, the resulting summary's semantic similarity to the source document is calculated. The outcomes of the experiments support the proposed model's success.

Text summarising techniques was introduced by Joshi [5] come in two different varieties. Includes extractive and abstractive summary. In this study, Summary of general extracted text is done using auto-encoders. The query-focused summarization, in contrast to the general summarization, demonstrates the document's key contents in accordance there with user-provided queries. Before sending each document to the sentence encoder's next stage, we first pre-processed it. Second, the objective of an encoder network is to transform an English sentence into a vector, which is then attempted to be translated into nearby sentences by the decoder. The connection between a source and a destination is made possible by the encoder-decoder design, which forces the model to pick out and abstract some crucial aspects. In order to map phrases with comparable semantic and syntactic properties into identical vector representations, the entire network is trained to reconstruct the surrounding sentences. An

auto-encoder is a specific kind of feed-forwarded neurons system developed to minimize data dimensionality. The input and output layers are identical, along with at least single secret layer with smaller dimensions than the input data is included. We only maintained the encoder portion of the auto-encoder network after training it, discarding the decoder portion, to produce a lower dimensional representation of each textual-unit embedding. The lower-dimension encoder output was referred to as "Latent Representation." This hidden representation of the entire text is merely a synopsis of it.

In [6] the digital world, as the amount of data produced at every instance is very huge .There is an ultimate need to develop a machine that can reduce the length of the texts automatically. Applying text summarization gears up the procedure of researching reduce reading time and increase the amount of information generated in the specific field. In this project, they tried to create a model as the solution which is based on extractive approach for summarization text, starting with NLP as the fundamental model. The extractive approach is actually successful in delivering the summery using the set of words which is actually most improve words in the actual text, hence the relevant information in figure 4. Existing model were created based on NLTK that is a library for processing text string by string. The input and output using NLTK[7] is the sequence of characters that is string. Provide various algorithms for a particular problem is the time consuming. The proposed model utilizes the library Spacy which selects the best option itself becoming more efficient.

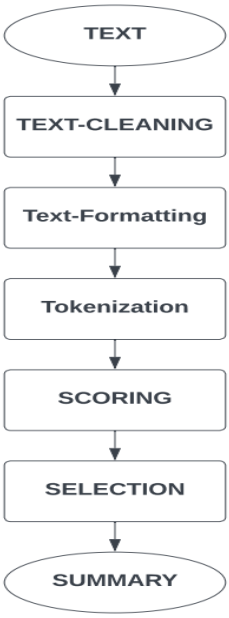

Figure 4. Step-by-step Implementation

3. Comparative Study:

The table below offers a brief comparison of the various approaches in NLP used in the summary process in order to summarise the text using adaptive models. This is a summary analysis that was produced from research that was mostly focused on NLP approaches and was published in a variety of magazines[8] in Table 2.

Table 2. Comparison of different Techniques used in Summarization of the Text.

Ref.No	Author	Approach	ROUGE SCORE		
			Rouge-1	Rouge-2	Rouge-L
[1]	Jiawen Jiang, Haiyang Zhang, Chenxu Dai, Qingjuan Zhao, Hao feng, Zhanlin Ji and Ivan Ganchev	ESN 'DA-PN' 'DA-PN + Cover' 'DA-PN + Cover + MLO'	**29.17** **32.20** **33.43** **33.75**	**10.97** **13.46** **13.77** **14.09**	**30.95** **31.26** **32.93** **32.69**
[2]	Nabil Alami, Mohammed Meknassi and Noureddine En-nahnahi	Variational auto encoder model, Auto-Encoder (AE), Extreme Learning Machine Auto-Encoder model (ELM-AE)	**13.34** **14.92** **16.37**	**25.91** **28.12** **26.71**	- - -
[3]	Shuai Zhao, Fucheng You and And Zeng Yuan Liu	BSA BSA* (Encoder (BERT) and Decoder (Transformer))	**40.4** **44.2**	**25.9** **28.9**	**36.7** **39.2**
[4]	Fucheng You, Shuai Zhao and And Jingjing Chen	TIF-SR (On LCSTS dataset) (On NLPCC2017 Dataset)	**42.1** **38.9**	**28.6** **24.6**	**37.4** **33.7**
[5]	Akanksha J, Alegre E, Fidalgo E and Laura Fernández-Robles	SummCoder (On AE-Net1) (On DUC 2002 Dataset) (On TIDSumm Dataset) (On Blog Summarization Dataset)	**51.7** **51.7** **58.8** **78.0**	**27.5** **27.5** **48.9** **71.7**	**44.6** **44.6** **49.3** **72.7**
[6]	Swaranjali Jugran, Ashish Kumar, Bhupendra Singh Tyagi and Mr. Vivek Anand	SpaCy & NLTK	-	-	-

4. Conclusion

NLP[10] is a field that must always advance in keeping with contemporary technological advancements. The summarising sector will greatly benefit from the incorporation of these techniques, both for consumers and developers. To develop more workable answers to the

issues faced by many readers, additional research must be conducted in this area. The future scope of this work will be to answer the real-time issues that are frequently experienced by users and need to be solved utilising the advanced methodologies. This study makes two significant contributions. The initial step entails conducting a thorough analysis of research articles that use Deep Learning [9][11]and approaches in NLP. Each study elaborates on the innovation, models, and experiments. The second step involves looking at the methods for developing text summarization models. Other text summarization techniques, like SpaCy, NLTK, Summy, and Gensim, can be examined.

Refrences

[1]. Jiang, J., Zhang, H., Dai, C., Zhao, Q., Feng, H., Ji, Z., & Ganchev, I. (2021). Enhancements of attention-based bidirectional lstm for hybrid automatic text summarization. IEEE Access, 9, 123660-123671.

[2]. Alami, N., Meknassi, M., & En-nahnahi, N. (2019). Enhancing unsupervised neural networks based text summarization with word embedding and ensemble learning. Expert systems with applications, 123, 195-211.

[3]. Zhao, S., You, F., & Liu, Z. Y. (2020). Leveraging Pre-Trained language model for summary generation on short text. IEEE Access, 8, 228798-228803.

[4]. You, F., Zhao, S., & Chen, J. (2020). A topic information fusion and semantic relevance for text summarization. IEEE Access, 8, 178946-178953.

[5]. Joshi, A., Fidalgo, E., Alegre, E., & Fernández-Robles, L. (2019). SummCoder: An unsupervised framework for extractive text summarization based on deep auto-encoders. Expert Systems with Applications, 129, 200-215.

[6]. JUGRAN, S., KUMAR, A., TYAGI, B. S., & ANAND, V. (2021, March). Extractive automatic text summarization using SpaCy in Python & NLP. In 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 582-585). IEEE.

[7]. Li, Z., Peng, Z., Tang, S., Zhang, C., & Ma, H. (2020). Text summarization method based on double attention pointer network. IEEE Access, 8, 11279-11288.

[8]. Adhikari, S. (2020, March). Nlp based machine learning approaches for text summarization. In 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC) (pp. 535-538). IEEE.

[9]. Afsharizadeh, M., Ebrahimpour-Komleh, H., & Bagheri, A. (2018, April). Query-oriented text summarization using sentence extraction technique. In 2018 4th international conference on web research (ICWR) (pp. 128-132). IEEE.

[10]. Irani, D., Webb, S., Pu, C., & Li, K. (2010). Study of trend-stuffing on twitter through text classification. In Collaboration, Electronic messaging, Anti-Abuse and Spam Conference (CEAS).

[11]. Aguru, A. D., Babu, E. S., Nayak, S. R., Sethy, A., & Verma, A. (2022). Integrated Industrial Reference Architecture for Smart Healthcare in Internet of Things: A Systematic Investigation. Algorithms, 15(9), 309.

[12]. Sethy, A., Rout, A. K., & Nayak, S. R. (2022, January). Face Recognition Based Automated Recognition System. In 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 543-547). IEEE.
[13]. Sarangi, P. K., Sahoo, A. K., Nayak, S. R., Agarwal, A., & Sethy, A. (2022). Recognition of Isolated Handwritten Gurumukhi Numerals Using Hopfield Neural Network. In Computational Intelligence in Pattern Recognition (pp. 597-605). Springer, Singapore.
[14]. Sethy, A., & Patra, P. K. (2021, March). Discrete Cosine Transformation Based Approach for Offline Handwritten Character and Numeral Recognition. In Journal of Physics: Conference Series (Vol. 1770, No. 1, p. 012004). IOP Publishing.