

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

Meeting Transcriber

Prof. Sharmila Chidaravalli, Tejas K S, Yadu Nandan B

Department of Information Science and Engineering Global Academy of Technology, Bangalore, India c.sharmila11@gmail.com, tejasks1ga20is144@gmail.com, yadu1ga20is161@gmail.com

Abstract: In a world where collaboration is increasingly digital, the efficiency and security of meetings have become paramount. Our project introduces a comprehensive online meeting platform extension that leverages advanced technologies to enhance the meeting experience. Using BERT for data classification, question answering, and summarization, along with an advanced Language Model (LLM) for data protection, we address common challenges such as information overload, language barriers, limited searchability, and a lack of actionable insights. This extension aims to revolutionize meetings, making them more productive, collaborative, and secure

Keywords: BERT (Bidirectional Encoder Representations from Transformers,), Text Classification, Key-Based Searching, LLM (Large language models)

I. INTRODUCTION

Welcome to the future of meetings! Our project is all about making online meetings better. Imagine meetings where it's easy to understand, find information, and keep things private.

Meetings can be tricky:

- Too much information: There's a ton of stuff discussed, and it's hard to remember what's important.
- Different languages: Sometimes, teams speak different languages, making it tough to understand each other.
- Hard to find things: Imagine searching for something specific in a really long meeting recording not fun!
- Not enough insights: Getting useful insights from discussions can be tricky without good tools.
- But guess what? We've got cool tech, like BERT, to help us make sense of data and answer questions. And
 we're adding an extra layer of protection called Language Model (LLM) to keep things super private and
 secure.
- Our dream is to turn meetings into something more than just talking. We're focusing on easy recording, summarizing, and finding stuff quickly. Plus, our top priority is keeping your info safe with the LLM.

In section 2, it consists of the brief literature review. In section 3, it consists of table of comparison results. section 4 provides conclusion and references

II. LITRATURE REVIEW

Analyses and references to real studies on NLP and AIML are included in the literature review

This study [1] discuss the methodology for the survey of the most recent neural abstractive text summarization models includes a thorough analysis of pertinent scientific literature architecture of encoder-decoder, mechanism of attention, Reinforcement learning-based training techniques, Distributional semantic rewards for training. Datasets from CNN/Daily Mail, DUC 2004, Giga word, and Newsroom. The research accomplished a thorough analysis of the most advanced neural abstractive text summarization models available today. It gave information about the methods, training plans, datasets, assessment measures, and design components of abstractive summarization models. The study also emphasized the difficulties and problems with these systems, which improved knowledge of the area and opened the door for more investigation and advancement in automatic text summarization.

The paper [2] focuses on Text summarizing strategies using several Large Language Models (LLMs) are examined in the article "Text Summarization Using Large Language Models: A Comparative Study of MPT-7b-instruct, Falcon-7b-instruct, and OpenAI Chat-GPT Models" by Lochan Basya, Mihir Sanghvi, et al. The methodology consists of studies

DOI: 10.48175/IJARSCT-15356

Copyright to IJARSCT www.ijarsct.co.in

2581-9429



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

using supervised and unsupervised summarizing techniques carried out on two different datasets: CNN/Daily Mail 3.0.0 and XSum. The study offers a thorough investigation of text summarization methods with LLMs, including the OpenAI Chat-GPT, Falcon-7b-instruct, and MPT-7b-instruct models.

The paper [3] focuses on a two-stage decoder framework for Multi-Document Scientific Summarization (MDSS) is presented in the paper "Multi-Document Scientific Summarization from a Knowledge Graph-Centric View," written by Pancheng Wang, Shasha Li, Kunyuan Pang, Liang liang He, Dong Li, Jintao Tang, Ting Wang, et al. in 2023. To improve content and connection modelling, KGSum, the suggested approach, uses knowledge graphs created from input scientific papers. Graph Updater, Entity-Sentence Updater, KGtext Generator, Summary Generator, copy mechanism, and combined training with an auxiliary decoder for knowledge graph representation and model stability are some of the components of the methodology. By outperforming other models on the multi-Xscience dataset, a large-scale MDSS dataset, and producing better outcomes in the domain, the study illustrates the efficacy of KGSum.

Thestudy [4] discusses using pre-trained language models (BERT, ALBERT), sequence-to-sequence (Seq2Seq) models, attention mechanisms, transformers, keyword extraction, sentence division, data preprocessing, and model fine-tuning, the paper presents a novel method for short text summarization based on keyword templates. As part of the methodology, preprocessing data is done by extracting keywords and classifying text according to these keywords. Transformers are integrated for decoding and BERT for encoding in the model architecture. The study demonstrates enhanced performance over baseline models and higher ROUGE scores on the LCSTS (Large Scale Chinese Short Text Summarization) dataset. Among the contributions are a customized model architecture and an altered data pretreatment technique intended for Chinese short text summarizing.

This paper [5] reveals that by modifying event cues, this multi-granularity summary system—which comprises of an Event Selector and an Event-aware Summarizer—achieves adaptable summarization. The need of multi-granularity summarization systems in practical situations is emphasized in the work, along with the necessity of benchmarks. The authors demonstrate the efficacy of the suggested GRANUSUM framework for unsupervised multi-granularity summarization by introducing their own dataset, GranuDUC, and conducting tests onmulti-news, arXiv, DUC2004, and GranuDUC.

The study [6] given presents the Dual-Attention Pointer Network (DAPT) model, an enhanced model for abstractive text summarization. Bidirectional LSTM encoders, self-attention mechanisms, soft attention mechanisms, a pointer structure, and an enhanced coverage method to handle problems like recurrence in generated summaries are some of the important elements incorporated into the architecture. By addressing problems with attention, coverage, and training objectives, the proposed DAPT model seeks to enhance previous models and offer a complete solution for abstractive text summarizing tasks. Through ROUGE-based assessments on the CNN/Daily Mail Dataset and the LCSTS Dataset, the experiments compare the model's efficacy with state-of-the-art models on various datasets. In order to provide accurate and efficient text summarization, this research presents the dual-attention pointer network (DAPT), which combines gate mechanisms and self-attention.

The study [7] presents a 10-step sequential technique that incorporates subject modelling, transformer-based punctuation restoration, and embedding-based sentence selection for extractive summarization of call transcripts. Effective tabulation and summary production are guaranteed by this process. Large Language Models (LLMs), partitional clustering, and internal validation indices are all incorporated into the approach; nevertheless, there are recommendations for enhancement that include adding LLMs, growing the number of terminals, and taking outside data into account. The efficacy of the extractive summarization technique is demonstrated by the studies conducted on the CNN/Dailymail and New York Times datasets. This method combines sentence selection, subject modelling, channel separation, and punctuation restoration to address issues with call transcript summarization. The paper advances the field by introducing a new punctuation restoration accuracy metric and a novel BERT transformer-based model [7].

Copyright to IJARSCT DOI: 10.48175/IJARSCT-15356 2581-9429 IJARSCT UJARSCT 410 www.ijarsct.co.in



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

information. Using genetic algorithms and genetic programming, the work suggests automatically generating an objective function for unsupervised text summary tasks using the DUC02 and CNN/Daily Mail datasets. It is discovered that the finest outcomes come from combining lexical and semantic information in the ideal way, advancing the field of extractive text summarization.

The research [9] focuses on the abstractive summarizing of texts consisting of a single document is the main emphasis of the paper's novel approach to text summarization. To improve the quality of generated summaries, the suggested TIF-SR (subject Information Fusion and Semantic Relevance) model computes semantic relevance and includes subject keyword information. Adjusting BERT, Attention Mechanism, Transformer Networks, and Semantic Similarity Calculation are all part of the process. The tests, carried out on the LCSTS and NLPCC2017 datasets, reveal that, in comparison to baseline models, the summaries produced by the TIF-SR model are more topical, have greater semantic similarity with the original content, and have better fluency and sentence coherence. The study offers a thorough examination of the test findings, emphasizing how well topic information and semantic relevance computation are integrated into the text summarizing task.

The study's [10] methodology for recommending tailored suggesting customized summaries of instructional materials is described in the paper. It makes use of formative assessment in the form of multiple-choice examinations, text preparation, and summary to gauge students' comprehension. In order to maximize summarizer performance, the configuration of the algorithm is researched, and learner-generated data from actual learning contexts—specifically, textbook data—are used to quantitatively assess summarizer performance. The suggested approach examines lengthy learning materials and generates brief text summaries based on the results of multiple-choice exams. A university-level B.S. course used the methodology to test it, and the results showed that it was comparable to teacher recommendations. A unique code summarization approach called Fret is presented in the paper[11] and is based on a reinforcertransformer architecture. With the goal of tackling issues like code comprehension and extended dependencies, Fret is made to produce functionally clear, accurate, and thorough code descriptions while delivering exceptional performance. Three primary parts make up the model: the Decoder, Code Encoder, and Functional Reinforcer. A range of machine learning algorithms are used, such as training and inference, decoding, functional reinforcement, and code encoding. These methods make use of feed-forward neural networks, positional encoding, feed-forward networks, BertC, BertNL, Hadamard & Norm, and multi-head attention. The research makes use of Python and Java datasets that were gathered from GitHub and include documentation comments and snippets of Java code, with distribution tendencies that are comparable to those of the Java dataset [11].

HunEmBERT, an improved BERT model for categorizing sentiment and emotion in political communication, is presented in the work [12]. The transformers library's Trainer API is used to fine-tune the huBERT model's parameters for sentiment and emotion categorization. Evaluation techniques include manual review of sentences that were incorrectly classified and standard metrics like Precision and Recall for error analysis using confusion matrices. The BERT Model, Transformer-based Pipeline, fine-tuning methods, Trainer API, and a number of assessment measures are all used in the study. The ISEAR Dataset and the HunEmPoli Corpus are used in the experiments. In this study, a domain-specific corpus (HunEmPoli) for sentiment and emotion analysis in political writings written in Hungarian is being created. This corpus is then used to fine-tune the huBERT model for sentiment and emotion classification tasks. The assessment of the performance of the adjusted models shows state-of-the-art.

A reinforced abstractive text summarization model with semantic added reward is presented in the work [13]. The single-layer bi-directional LSTM encoder and decoder in the model is used in conjunction with a sequence-to-sequence attention mechanism. Interestingly, it uses intra-decoder attention to reduce problems such as duplicating hard words and generating repetitive phrases. In order to maximize metrics, the approach uses policy learning and reinforcement learning, with an emphasis on word similarity, semantic similarity, and word mover distance to guarantee readability. The Gigaword summarization dataset is used in the study. We offer two reward functions that augment n-gram matching with semantic values: ROUGE-SIM and ROUGE-WMD. Sequence-to-sequence, Transformer-based prelearning, and reinforcement learning models are all outperformed by the models created in this work. Readability and grammatical accuracy show significant gains, demonstrating the efficacy of the suggested strategy.

A innovative multi-step methodology for automatic news text summarizing is presented in the publication [14]. Preprocessing, segmentation, tokenization, stop-word removal, stem extraction, and the extraction of important features

Copyright to IJARSCT DOI: 10.48175/IJARSCT-15356

2581-9429

IJARSCT



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

at the word and sentence levels are all included in the process. To extract word features, the suggested approach makes use of a fuzzy logic system, a SpeGenetic Algorithm, and a particular keyword extraction technique. The DUC2002 datasets are used in the investigation. For news text summarizing, the model integrates fuzzy logic, evolutionary algorithms, and multi-feature. It extracts significant aspects, ranked according to the qualities of the news text, including word and sentence features. The fair weights are produced by use of the fuzzy logic system. The technique makes use of ROUGE-2 for the ideal distribution of weight and ROUGE-1 for fitness values. Results from experiments show that this approach works better than

A succinct and effective BERT-based methodology for identifying rumors on Twitter, called CE-BERT, is presented in the article [15]. The process consists of layer selection that works, efficient fine-tuning, and use of classification techniques. A number of layer reduction techniques are examined, with a focus on the significance of the fine-tuning procedure. These include the basic model, six-layer models, four-layer models, and two-layer models. The Twitter 15 Dataset, Twitter 16 Dataset, and PHEME Dataset are used in the study. The study shows that CE-BERT is a low-computing-requirement effective model for Twitter rumor identification. It performs better than the state-of-the-art models, especially when it comes to source text scenarios, demonstrating its efficacy in Twitter rumor identification. This study [16] discusses the methodology for Korean abstractive text summarization utilizing a Multi-Encoder

This study [16] discusses the methodology for Korean abstractive text summarization utilizing a Multi-Encoder Transformer model. Key components include Transformer-based Encoder-Decoder Model, Multi-Encoder Architecture, Combining Representations, Fine-Tuning with BERT-based PLMs, Auto-regressive Decoding, and Evaluation Metrics such as ROUGE and BERTScore. Datasets utilized include Law (AI-Hub) dataset, News (AI-Hub) dataset, and News (NIKL) dataset. The paper introduces a transformer-based encoder-decoder model tailored for Korean abstractive summarization and proposes a multi-encoder architecture utilizing various pre-trained models. It demonstrates significant performance enhancements over single-encoder models, particularly with diverse encoders, across three Korean summarization datasets, while also conducting thorough ablation studies and qualitative analyses to assess the efficacy of different techniques, including OOV token handling.

[17] presents a text summarization method called Topic Information Fusion and Semantic Relevance (TIF-SR), which aims to enhance the quality of generated summaries by combining topic information and semantic relevance. Key components include BERT-based encoding, Topic Information Fusion, Transformer architecture, semantic relevance calculation, and fine-tuning objective. Datasets utilized include LCSTS and NLPCC2017 datasets. The TIF-SR model significantly improves ROUGE scores on LCSTS and NLPCC2017 datasets by incorporating topic information and semantic relevance, resulting in fluent, readable, and semantically relevant summaries.

[18] introduces a Topic Modeling-Based Framework for Extracting Marketing Information From E-Commerce Reviews. The methodology involves extracting strengths and weaknesses of individual products, identifying connections among different products, and predicting product trends. Techniques employed include clustering algorithms, topic modeling (such as BERTopic and DCN), and a Transformer-based forecasting model. General datasets are considered for the survey paper. The paper achieved significant advancements in extracting product characteristics, identifying related products, and forecasting product trends using topic modeling techniques and a Transformer-based model. Moreover, it provided valuable insights into consumer preferences and market trends through thorough analysis and evaluation.

[19] proposes a framework for automatic context extraction and comparison of short text documents, specifically focusing on service provider policies, with a particular emphasis on GDPR compliance and comparison with 3000 web service privacy policies. The framework includes six phases. Techniques employed include Entity Recognition, Modal Logic, Machine Learning, BERT Text Summarization, and Semantic Web Languages to extract high-frequency entities from GDPR documents. The dataset comprises over 3000 web service provider privacy policies, categorized into EU-based and worldwide policies, downloaded post-May 2018 to ensure relevance post-GDPR. The research utilizes NLP, deep learning, and semantic web techniques to efficiently analyze short or incomplete texts, focusing on privacy policies in progress.

[20] presents a Chinese text summary model leveraging keyword templates and fine-tuning with pretrained language models like BERT. The methodology involves data preprocessing, model architecture design, and experimentation. The study explores pretrained language models, extractive models, and abstractive models for language understanding tasks, extractive summarization, and abstractive summarization, incorporating attention mechanisms and BERT. The LCSTS

Copyright to IJARSCT DOI: 10.48175/IJARSCT-15356 2581-9429 IJARSCT UJARSCT 412 www.ijarsct.co.in



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

dataset is utilized. The proposed BSA model outperforms baseline models on the LCSTS dataset, enhancing ROUGE scores and text division quality. Future work should explore its applicability to other pretrained language models.

The text discusses a proposed abstractive summarization model [21] that builds upon the BART architecture by adding a discriminator and a cluster generator. The model splits original text into sentences and classifies each sentence as salient or non-salient, creating clusters for salient and non-salient sentences. Experiments on XSUM and CNN/DailyMail datasets showed the model outperformed BART and PEGASUS in terms of ROUGE and BERTScore. Various techniques for abstractive summarization are utilized, including the Encoder-Decoder Attention Model, Pointer Network, Coverage Mechanism, Transformer Model, UniLM, BERTSUM and BART Models, PEGASUS Model, Reinforcement Learning, Topic Model, Multimodal Information, Attention Head Masking, Information Theory, Extraction-and-Paraphrasing, Entity Aggregation and Factuality Consistency, Deep Communicating Agents, Sentence Correspondence, Graph, and Bottom-Up Approach. The study uses CNN/DailyMail Dataset and XSUM Dataset. A new pre-trained sequence-to-sequence model improves abstractive summarization by extracting salient sentences and clustering context vectors. Experimental results outperform existing models, addressing fact inconsistency and long document summarization.

The algorithm proposed by Yu et al. [22] utilizes a pre-trained BERT language model to vectorize the sentence-level features of policy text, followed by classification using a classifier. BERT (Bidirectional Encoder Representations from Transformers), a transformer-based language model pre-trained on large corpora of text data, is employed in this technique. The paper uses policy text data from various policy domains. The trained model achieves an improvement in policy text classification accuracy, with the highest F1 value of 93.25% on the test set. This result demonstrates a nearly 6% increase in performance compared to the BERT model's classification task for the MRPC (Microsoft Research Paraphrase Corpus) task.

The research conducted by Liu et al. [23] focuses on text classification, combining the Bert model with Bayesian networks. Initially, the Bayesian network is employed for coarse classification into two categories, providing an approximate category range for each text. Subsequently, the Bert model is utilized for fine-grained classification within the identified range. The techniques employed include the Bert model, a pre-training model based on deep learning for natural language processing (NLP), and Bayesian networks, a probabilistic graphical model representing a set of random variables and their conditional dependencies via a directed acyclic graph. The paper utilizes text data related to people's livelihood governance, containing unstructured data from various sources within this domain. By combining the Bayesian network's coarse classification with the Bert model's fine-grained classification, the paper achieves an improvement in text classification accuracy. This approach reduces errors caused by classification defects inherent in using only one of the methods, ultimately enhancing the overall accuracy of text classification.

The research conducted by Seo et al. [24] focuses on classifying and interpreting unstructured traffic crash description data through three stages: data preprocessing, model development, and model performance evaluation. Techniques utilized include word embedding algorithms, a BERT-based text classification model, and cross-validation for robustness. The study analyzes 2,427 traffic crash datasets from the Korean Traffic Crash Analysis System in Daegu, Korea, focusing on crash data elements, weather conditions, and written descriptions. The developed method classifies traffic crash text into standardized data using the BERT model, achieving a text classification accuracy of over 95% for most labels. The study aims to increase the utilization of underutilized traffic crash data in traffic safety research, ultimately reducing crashes and improving traffic safety.

The research by Tian and Wang et al. [25] introduces a scientific document retrieval system comprising three modules: Document Preprocess, Mathematical Expression Similarity, and Text Similarity. These modules are responsible for extracting mathematical expressions, unifying LaTeX and MathML, and calculating similarity between text parts. Techniques employed include FDS for decomposition, HFS for calculating similarity, BERT for text similarity, keyword extraction, and cosine similarity to analyze mathematical expressions. The dataset consists of scientific documents containing mathematical expressions and related texts, from which keywords are extracted to train BERT and calculate text similarity. The study proposes a retrieval method that combines mathematical expressions and contextual text for mathematical information retrieval, incorporating membership degree, HFS, BERT, context keywords, and word vectors. Future research aims to improve extraction methods, address pressions and evaluate Chinese and English algorithms.

DOI: 10.48175/IJARSCT-15356

Copyright to IJARSCT www.ijarsct.co.in

413

2581-9429

IJARSCT



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

An et al. [26] present a Topic Modeling-Based Framework for Extracting Marketing Information From E-Commerce Reviews, automating product extraction, pros and cons identification, and trend forecasting from Naver Shopping review data using clustering algorithms and a Transformer-based forecasting model. Techniques utilized include Mecab, POS analysis, KcBERT, topic modeling, and co-occurrence frequency to identify related products. The study employed clustering algorithms, dimension reduction algorithms, topic modeling, text summarization, and transformer-based forecasting to analyze review data, extract insights, and predict trends. Analyzing 530,877 consumer reviews from Naver Shopping for 17 products, the study identified key POS terms and analyzed search volume trends using KcBERT, TF-IDF, and clustering algorithms. Utilizing topic modeling to extract product characteristics and related products from consumer reviews, the study aimed to uncover marketing insights. Although the DCN-based model performed well, its results were slightly less favorable than the baseline BERTopic. The insights gleaned can guide marketing strategies, identify products with similar intentions, and inform product introductions. The product trend prediction model also improved its predictive performance by extracting key product information from text data, though it cannot predict future changes due to identifying past data patterns.

Wasim et al. [27] introduce a Multi-Label Question Classification framework for factoid and list type questions in biomedical question answering. The methodology comprises two main components: Multi-label Biomedical Question Classification and Proposed Biomedical Question Answering Methodology. Preprocessing involves tokenization, lemmatization, and dependency parsing, followed by feature extraction, data transformation, classification, and collective re-ranking. The study utilizes the MLBioMedLAT corpus and BioASQ dataset. The paper presents a novel methodology for classifying biomedical questions into multiple labels, achieving enhanced accuracy through sophisticated preprocessing and classification techniques. Additionally, the proposed question answering approach demonstrates superior performance compared to existing systems, surpassing baseline metrics and showcasing its efficacy in biomedical question answering tasks.

Meng et al. [28] present a study focusing on efficiently classifying electric power audit texts using pre-training and fine-tuning models such as BERT, MAE, and CLIP. These models require less supervised data for fine-tuning and outperform non-pre-trained models. The EPAT-BERT model introduces two pre-training tasks specifically tailored for electric power audit text classification using BERT's structure: word-granularity masked language modeling and entity-granularity masked language modeling, which enhance the understanding of electric power audit texts. The study utilizes a dataset comprising 1.5M electric power-related texts and 1,500 audit texts, alongside baseline models including Naive Bayes, SVM, GBDT, AdaBoost, XGBoost, TextCNN, LSTM, and BERT. The EPAT-BERT model demonstrates effectiveness in classifying electric power audit texts through two ablation experiments. Results indicate that the model significantly outperforms existing models in accuracy, precision, recall, and F1 score compared to traditional machine learning models. The model's extensibility allows for its application to other electric power text classification tasks and project type annotation.

El-Kassas et al. [29] conducted a comprehensive survey of Automatic Text Summarization (ATS) techniques and methodologies. The methodology involved discussing various approaches, methods, building blocks, techniques, datasets, evaluation methods, and future research directions in ATS. The paper covers extractive, abstractive, and hybrid approaches to ATS, highlighting their differences and applications. Additionally, it explores different building blocks and techniques used in ATS, along with the datasets commonly employed for training and evaluation purposes. Although the survey paper does not contain any datasets, it achieves the goal of providing researchers with a comprehensive overview of ATS, serving as a valuable resource for understanding the current state of the field, identifying research gaps, and guiding future directions for improving ATS techniques.

Table of Comparison Result:

	Tuole of Companion result.							
Sl	Author/	Research /Work	Methodology	Technique	Dataset/Input	Experiment/		
no	year	Paper				Observation		
1	AYESHA	A Survey of the	The	Encoder-	CNN/Daily	The paper achieved a		
	AYUB	State-of-the-Art	methodology	decoder	Mail dataset,	comprehensive review		
	SYED,	Models in	used in the	architecture,	DUC 2004	of the current state-of-		
	FORD	Neural	survey of state-	Attention	dataset,	the-art models in neural		
	LUMBAN	Abstractive Text	of-the-art models	mechanism,	Gigaword	abstractive text		

Copyright to IJARSCT



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 1, February 2024

	GAOL,T	Summarization	in neural	Training	dataset,	summarization. It
	OKURO		abstractive text	methods	Newsroom	provided insights into
	MATSUO		summarization	based on	dataset	the design elements,
	et al.		involves a	reinforcement		mechanisms, training
	(2020)		comprehensive	learning,		strategies, datasets, and
	`		review of	Distributional		evaluation metrics used
			relevant	semantic		in abstractive
			scientific	rewards for		summarization models.
			literature	training		Additionally, the paper
				S		highlighted the
						challenges and issues
						associated with these
						systems, contributing to
						a better understanding
						of the field and paving
						the way for future
						research and innovation
						in automatic text
						summarization
2	Lochan	Text	Methodology	The	CNN/Daily	This research embarked
	Basya	Summarization	used in this study	techniques	Mail 3.0.0	on a comprehensive
	Mihir	Using Large	involves	used in the	and the	exploration
	Sanghvi	Language	conducting	study include	Extreme	of text summarization
	et al.	Models:	experiments on	supervised	Summarizatio	techniques using
	(2022)	A Comparative	various Large	and	n (XSum)	various Large Lan-
	(')	Study of MPT-	Language	unsupervised	dataset.	guage Models (LLMs
		7b-instruct,	Models (LLMs)	summarizatio		86.
		Falcon-7b-	using two	n methods.		
		instruct, and	distinct datasets,			
		OpenAI Chat-	CNN/Daily Mail			
		GPT Models	3.0.0 and XSum			
		Lochan Basyal				
3	Pancheng	Multi-Document	The	he model uses	The Multi-	The study proposes a
	Wang,	Scientific	methodology	Graph	Xscience	knowledge graph-
	Shasha Li,	Summarization	described in the	Updater,	dataset, a	centric Transformation-
	Kunyuan	from a		Entity-	large-scale	based model for
	Pang,	Knowledge	involves a two-	Sentence	MDSS	MDSS, which
	Liangliang	Graph-Centric	stage decoder	Updater,	dataset	significantly
	He, Dong	View	framework for	KGtext		outperforms other
	Li, Jintao		Multi-Document	Generator,		models and achieves
	Tang,		Scientific	Summary		the best results on the
	Ting,		Summarization	Generator,		multi-Xscience dataset.
	Wang. et		(MDSS). The	copy		
	al.		model, named	mechanism,		
	(2023)		KGSum, utilizes	and joint		
	()		knowledge	training with		
			graphs	an auxiliary		
			constructed from	decoder for		REPARCH IN SOUTH
						ISSN

Copyright to IJARSCT www.ijarsct.co.in



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

			input scientific	knowledge		
			papers to	graph		
			enhance content	representation		
			and relationship	and model		
			modeling.	stability.		
4	SHUAI	Leveraging Pre-	The paper	Pre-trained	LCSTS	Proposes a novel
	ZHAO,	Trained	proposes a novel	Language	(Large Scale	approach for short text
	FUCHEN	Language Model	approach for	Models	Chinese Short	summarization using
	G YOU,	for Summary	short text	(BERT,	Text	pre-trained language
	AND	Generation on	summary	ALBERT),	Summarizatio	models and keyword
	ZENG	Short Text	generation based	Sequence-to-	n) dataset	templates.
	YUAN		on keyword	Sequence	n) dataset	Introduces a modified
	LIU et al.		templates. This	(Seq2Seq)		data preprocessing
	(2022)		approach	Models,		method and model
	(2022)		involves	Attention		architecture tailored for
			preprocessing	Mechanism,		Chinese short text
			data by	Transformers,		summarization.
			extracting	Keyword		Demonstrates improved
			keywords and	Extraction,		performance compared
			dividing text	Sentence		to baseline models, as
			based on these	Division,		evidenced by
			keywords. The	Data		higher ROUGE scores.
			paper describes	Preprocessing		8
			the model	, Model Fine-		
			architecture,	Tuning		
			which utilizes a	8		
			combination of			
			BERT for			
			encoding and			
			Transformers for			
			decoding.			
			Selection.			
5	Ming	Unsupervised	The text	The text	Multi-News,	This paper highlights
	Zhong,	Multi-	discusses a	discusses	arXiv,	the significance of
	Yang	Granularity	GRANUSUM	customized	DUC2004,	multigranularity
	Liu†Suyu	Summarization	multi-granularity	summarizatio	and the newly	
	Ge,Yunin		summarization	n,	introduced	in real-world scenarios,
	g Mao,		framework,	unsupervised	GranuDUC	proposing the first
	Yizhu		comprising an	approaches,		unsupervised multi-
	Jiao,Xing		Event-aware	and multi-		granularity
	xing		Summarizer and	granularity		summarization
	Zhang,		an Event	summarizatio		framework
	Yichong		Selector, which	n models,		GRANUSUM, and
	Xu,Cheng		achieves multi-	emphasizing		demonstrating its
	uang Zhu,		granularity	the need for		effectiveness through
	Michael		summarization	benchmarks		experiments
	Zeng,		by adjusting	and		
	Jiawei		event hints	introduces	ls.	RECARCH IN SOURCE
L		<u> </u>	<u>i</u>	<u> </u>		ISSN

Copyright to IJARSCT www.ijarsct.co.in



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 1, February 2024

	TT 0 . 1			.1 .		
	Han§et al.			their own		
	(2022)			dataset,		
				GranuDUC.		
6	Zhixin Li,	Text	The presented	The proposed	LCSTS	This paper introduces a
	Zhi Peng,	Summarization	work describes	DAPT model	Dataset,	dual-attention pointer
	Suqin	Method Based on	an improved	aims to	CNN/Daily	network (DAPT) for
	Tang,	Double Attention	model for	improve upon	Mail Dataset	text summarization,
	Canlong	Pointer Network	abstractive text	existing		combining self-
	Zhang,		summarization,	models by		attention and gate
	Huifang		referred to as the	addressing		mechanisms for
	Maet al.		Dual-Attention	issues related		accurate and consistent
	(2020)		Pointer Network	to attention,		summary generation,
	(====)		(DAPT) model.	coverage, and		improving ROUGE
			The architecture	training		evaluation index.
			incorporates	objectives,		
			several key	providing a		
			components,	comprehensiv		
			including	e solution for		
			bidirectional	abstractive		
			LSTM encoders,	text		
			self-attention	summarizatio		
			mechanisms, soft	n tasks. The		
			attention			
				experiments involve		
			mechanisms, a			
			pointer structure,	comparisons		
			and an improved	with state-of-		
			coverage	the-art		
			mechanism to	models on		
			address issues	different		
			such as repetition	datasets,		
			in generated	highlighting		
			summaries	the model's		
				effectiveness		
				through		
				ROUGE-		
				based		
				evaluations		
7	PRATIK	Extractive	The proposed	The method,	CNN/Dailym	This paper presents an
	K.	Summarization	10-step	incorporating	ail, New	extractive
	BISWAS	of Call	sequential	LLMs,	York Times	summarization
	1 AND	Transcripts	procedure for	partitional	datasets	technique that
	ALEKSA		extractive	clustering,		addresses challenges in
	NDR		summarization	and internal		call transcript
	IAKUBO		of call transcripts	validation		summarization by
	VICH et		includes topic	indices, is		combining channel
	al.		modeling,	proposed for		separation, topic
	(2022)		embedding-	improvement,		modeling, sentence
			based sentence	with	ls.	selection, and
	I	I	I			ISSN

DOI: 10.48175/IJARSCT-15356

Copyright to IJARSCT www.ijarsct.co.in

417



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

			selection, and	suggestions		punctuation restoration.
			transformer-	for		It uses a novel BERT
			based	incorporating		transformer-based
			punctuation	LLMs,		model and a new metric
			restoration,	expanding		for punctuation
			ensuring	terminals, and		restoration accuracy
			efficient	considering		
			summary	external		
			generation and	information		
			tabulation			
8	ÁNGEL	Toward the	The proposed	The method,	DUC02 and	The study proposes an
	HERNÁN	Automatic	method aims to	incorporating	CNN/Daily	automatic generation of
	DEZ-	Generation of an	create fitness	LLMs,	Mail datasets	an objective function
	CASTAÑ	Objective	functions for	partitional		for unsupervised text
	EDA,REN	Function for	evolutionary	clustering,		summary tasks using a
	É	Extractive Text	cluster-based	and internal		genetic algorithm and
	ARNULF	Summarization	methods for	validation		genetic programming.
	О		automatic text	indices, is		The optimal
	GARCÍA-		summarization.	proposed for		combination of lexical
	HERNÁN		It involves	improvement,		and semantic
	DEZ2		representing	with		information is found to
	AND		texts or	suggestions		yield the best results.
	YULIA		documents as	for		J
	LEDENE		numeric vectors	incorporating		
	VA et al.		using different	LLMs,		
	(2023)		mapping	expanding		
	(2023)		methods. Four	terminals, and		
			methods are	considering		
			proposed: term	external		
			frequency-	information.		
			inverse	miormation.		
			document			
			frequency (tf -			
			idf), one-hot			
			encoding (OHE),			
			and latent			
			Dirichlet			
			allocation			
			(LDA). These			
			methods explore			
			the relevance of			
			lexical and			
			semantic			
			information for			
			identifying			
			relevant			
			sentences in the			
			ATS task			SSEARCH IN SCIENCE



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 1, February 2024

0	FUGUENI		TO I	T	I como	
9	FUCHEN	A Topic	The	Fine-tuning	LCSTS,	Shows that the
	G YOU,	Information	methodology	BERT,	NLPCC2017	generated summaries
	SHUAI	Fusion and	used in the paper	Attention		by the proposed model
	ZHAO,	Semantic	involves	Mechanism,		are closer to the topic,
	AND	Relevance for	proposing a new	Transformer		have higher semantic
	JINGJING	Text	approach to text	Networks,		similarity with the
	CHEN et	Summarization	summarization,	Semantic		source document, better
	al.		focusing on	Similarity		fluency, and sentence
	(2020)		abstractive	Calculation		coherence compared to
	(2020)		summarization	Culculation		baseline models.
			of single-			Provides a
			document texts.			comprehensive analysis
			The paper			of experimental results,
			introduces the			highlighting the
			TIF-SR (Topic			effectiveness of
			Information			incorporating topic
			Fusion and			information and
			Semantic			semantic relevance
			Relevance)			calculation in text
			model, which			summarization task
			incorporates			
			topic keyword			
			information and			
			calculates			
			semantic			
			relevance to			
			improve the			
			quality of			
			generated			
			summaries.			
10	LUCA	Dagaman an din a		Famustina	#201 120##2#	This managements of
10		Recommending	The proposed	Formative	real learner-	This paper presents a
	CAGLIER	Personalized	methodology	Assessment,	generated	methodology for
	0,	Summaries of	uses formative	Text	data from	exploring large learning
	LAURA	Teaching	assessment with	Preparation,	textbooks	documents, providing
	FARINET	Materials	multiple-choice	he algorithm		short textual summaries
	TI,		tests, text	configuration		based on multiple-
	AND		preparation, and	is studied to		choice test outcomes.
	ELENA		summarization to	optimize		The method was tested
	BARALIS		evaluate learners'	summarizer		in a university-level
	et al.		understanding	performance.		B.S. course, showing
	(2019)			Quantitative		similarity to teacher
				evaluation of		recommendations.
				summarizer		
				performance		
				is conducted		
				on real		
				learner-		
				generated		ESEARCH IN SCIEN
				data in a real		ISSN
Conv	right to IJAR	SCT.	DOI: 10.48175/IJ			581-9429 419

Copyright to IJARSCT www.ijarsct.co.in



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

This paper introduces a introduces a introduces a language introduce and a language introduces and a language introduce and a concise and functionally clear, accurate, and functionally clear, accurate, and functionally clear, accurate, and descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. Reinforcer, Code understanding and long dependence, and achieving remarkable performance. Reinforcer, Code incoding, multi-head attention, feed-forward networks, BerrC, BertNL, and Hadamard & Norm, as well as feed-forward networks BERT Model, Transformer-language incomplete in the political components and incomplete in the political components in the language in the transformer and incomplete in the political components in the language in the langua		1				T	
Wang, Hanwen novel code Zhang, Guoliang Lu, Lei Lyu, and reinforcer- transformer arthitecture. Fret effectively generates functionally clear accurate, and descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVÁN UVEGES AND ORSOLY A RING ORSOLY A RING et al. (2021) Political Communication in cludded standard metrics library's Trainer API. Evaluation included standard metrics showing the transformers allusted using the transformers is library's Trainer API. Evaluation included standard metrics showing summer and emotion classification at transformer stending and elegandency and achieving remarkable performance.					learning context		
Wang, Hanwen novel code Zhang, Guoliang Lu, Lei Lu, Lei Lu, Lei Lyu, architecture. Fret effectively generates functionally clear accurate, and descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVÁN UVEGES AND ORSOLY A RING or RING at al. (2021) 12 ISTVÁN UVEGES AND ORSOLY A RING et al. (2021) 13 ISTVÁN UVEGES AND DERT-Model for ORSOLY A RING et al. (2021) 14 ISTVÁN UVEGES AND DERT-Model for ORSOLY A RING et al. (2021) 15 ISTVÁN UVEGES AND DERT-Model for ORSOLY A RING et al. (2021) 16 ISTVÁN Et al. (2021) 17 ISTVÁN UVEGES AND DERT-Model for ORSOLY A RING et al. (2021) 18 ISTVÁN UVEGES AND DERT-Model for ORSOLY A RING et al. (2021) 19 ISTVÁN Et al. (2021) 10 ISTVÁN UVEGES AND DERT-Model for ORSOLY A RING et al. (2021) 10 ISTVÁN ET A RING et al. (2021) 10 ISTVÁN ET A RING et al. (2021) 11 ISTVÁN ET A RING et al. (2021) 12 ISTVÁN ET A RING et al. (2021) 14 INDEMBERT: A IT A RING et al. (2021) 15 ISTVÁN ET A RING et al. (2021) 16 INDEMBERT A IT A RING et al. (2021) 17 INDEMBERT A IT A RING et al. (2021) 18 INDEMBERT A IT A RING et al. (2021) 19 INDEMBERT A IT A RING et al. (2021) 10 INDEMBERT A IT A RING et al. (2021) 10 INDEMBERT A IT A RING et al. (2021) 11 INDEMBERT A IT A RING et al. (2021) 12 ISTVÁN ET A RING et al. (2021) 13 INDEMBERT A IT A RING et al. (2021) 14 INDEMBERT A IT A RING et al. (2021) 15 INDEMBERT A IT A RING et al. (2021) 16 INDEMBERT A IT A RING et al. (2021) 17 INDEMBERT A RING et al. (2021) 18 INDEMBERT A RING et al. (2021) 19 INDEMBERT A RING et al. (2021) 10 INDEMBERT A RING et al. (2021) 10 INDEMBERT A RING et al. (2021) 10 INDEMBERT A RING et al. (2021) 11 INDEMBERT A RING et al. (2021) 12 ISTVÁN ET A RING et al. (2021) 13 INDEMBERT A RING et al. (2021) 14 INDEMBERT A RING et al. (2021) 15 INDEMBERT A RING et al. (2021) 16 INDEMBERT A RING et al. (2021) 17 INDEMBERT A RING et al. (2021) 18 INDEMBERT A RING et al. (2021) 19 INDEMBERT A RING et al. (2021) 19 INDEMBERT A RING et	11	Ruyun	This paper	The Fret model	various	The java and	This paper introduces a
Hanwen Zhang, Guoliang Lu, Lei Lyu, and Chen Lyu, and Chen Lyu et al. (2020) effectively generates functionally clear, accurate, and comprehensive code understanding and alloviating issues like code understanding and alloviating remarkable performance. 12 ISTVAN Hunembert: A The hubert performance. 13 ISTVAN BERT-Model for ORSOLY Classifying A RING et al. Emotion in (2021) Political Communication included standard metrics and emotion included standard metrics are desirable transformers adjusted using the transformers library's Trainer API. Evaluation included standard metrics and thuman-readable descriptions, alleviating issues like code understanding and a long dependency, and achieving remarkable performance. Hunembert: A The hubert included standard metrics show the fine fine fill the from Gill tub, contain Java reinforcement code encoding, dinctional training and inference. Code snippets and documentation and teal. Comprehensive code encoding, algorithms, including algorithms, including algorithms, collected from Gill tub, contain Java reinforcement code singulation and including and manuel training and inference. These and documentation and training and inference. These and documentation and training and inference. These and documentation and training and inference. These and documentation inference. These and documentation and training and inference. These and documentation and adacting training and inference. These and documentation and descriptions, alleviating issues like code understanding and numbers activities and adocumentation. These and documentation inference. These and documentation and descriptions and understanding and inference. These and documentation inference. These and documentation and descriptions algorithms inference. These and documentation inference. These and documentation and encoding, multi-head attention, feed-forward networks. BertC, BertNL, and Hadamard & Norm, as well as feed-forward networks. Trainsformers and the motion included standard metrics and the motion included		-	introduces a	is designed for	machine	Python	novel code
Zhang, Guoliang model called providing a Lyu, and Lyu, and Lyu, and Chen Lyu et al. (2020) Effectively generates functionally clear, accurate, and comprehensive code understanding and long dependency, and achieving remarkable performance. Emodel performance. Emodel standard metrics Emodel for Resifring and long dependency and achieving remarkable performance. Emodel standard metrics Example 1 Evaluation included Emodel standard metrics Evaluation of the fine-tuning the transformer and glisted using the transformer and classification task using this corpus Evaluation of the fine-tuning the consist of the surface and comprehensive code descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. East of the components the fine tuning and achieving remarkable performance. East of the components the fine tuning and long dependency, and achieving remarkable performance. East of the components the fine tuning and long dependency, and achieving remarkable performance. East of the contain Java reinforcer-transformer architecture. Free effectively generate functionally code and no comments, with similar inference. These algorithms use multi-head attention, positional encoding, multi-head att		-	novel code	-	learning	datasets,	summarization model
Guoliang Lu, Lei Lu, Lei Ly, and reinforcer- transformer description of its architecture. Free functionally clear, accurate, and descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVÁN TÜVEGES AND ORSOLY ARING CABLE AND ORSOLY CLASSIFYING A RING CABLE COMMUNICATION COMMUNICATION CASSIFICATION COMMUNICATION CASSIFICATION COMMUNICATION CASSIFICATION CASSIFICATION COMMUNICATION CASSIFICATION COMMUNICATION CASSIFICATION CASSIFICATION COMMUNICATION COMMUNICATION COMMUNICATION COMMUNICATION COCC de description dits reinforcer-transformer rediffectively generate code description of its reinforcement code description of its reinforcement code and functional components: components: the Functional comprehensive code code code code code code code cod		Zhang,	summarization	-	_	collected	called Fret, based on a
Lu, Lei Lyu, and Chen Lyu transformer reinforcer-transformer description of its architecture. Fret functionality. effectively generates functionality. clear, accurate, and comprehensive code understanding and long dependency, and achieving remarkable performance. 12 ISTVÁN TÜVEGES AND ORSOLY A RING ORSOLY A RING ORSOLY A RING Sentiment and et al. (2021) 12 ISTVÁN GERT-Model for ORSOLY A RING Sentiment and et al. (2021) 13 ISTVÁN GERT-MODEL STEIN COMMUNICATION ORSOLY A RING Sentiment and et al. (2021) 14 ISTVÁN COMMUNICATION ORSOLY A RING Sentiment and et al. (2021) 15 ISTVÁN COMMUNICATION ORSOLY A RING Sentiment and et al. (2021) 16 ISTVÁN COMMUNICATION ORSOLY A RING Sentiment and et al. (2021) 17 ISTVÁN COMMUNICATION ORSOLY A RING Sentiment and et al. (2021) 18 ISTVÁN COMMUNICATION ORSOLY A RING Sentiment and et al. (2021) 19 ISTVÁN COMMUNICATION ORSOLY A RING Sentiment and emotion included standard metrics of the transformers library's Trainer API. Evaluation included standard metrics of the condex of the functional functional training and inference code code encoding, encoding, encoding, decodeng, and training and inference. These algorithms decoding, and training and inference. These algorithms decoding, and training and inference distribution training and attention, positional encoding, multi-head attention, feed-forward networks, BertC, BertNI., and Hadamard & Norm, as well as feed-forward networks and provided descriptions, alleviating inference decoding, and training and inference. These algorithms decoding, and training and inference distribution training and dependency, and activation inference decoding, and training and decoding, and training and inference distribution training and attention, prositional encoding, and training and inference distribution training and attention, prositional encoding. The functio		-			_	from GitHub.	· · · · · · · · · · · · · · · · · · ·
Lyu, and Chen Lyu transformer transformer transformer transformer transformer transformer transformer architecture. Fret (2020) (2020) (2020) (2020) (2020) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2021) (2020) (2020) (2020) (2021		_			_		
Chen Lyu transformer et al. (2020) effectively effectively elear, accurate, and comprehensive code code understanding and long dependency, and achieving remarkable performance. 12 ISTVAN HunEmBERT: A TVS GNSOLY CLassifying AND BERT-Model for ORSOLY A RING et al. (2021) Political Emotion in (2021) Political Emotion in (2021) Political Emotion in (2021) Political Emotion in in (2021		*					
et al. (2020) effectively generates functionality. The model generates functionally clear, accurate, and comprehensive code descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVAN dependency, and achieving remarkable performance. 12 ISTVAN HunEmBERT: A Time hubErt Vergets AND BERT-Model for ORSOLY A RING Sentiment and et al. (2021) Political Communication in (2021) Political Communication in classification, with parameters library's Trainer API. Evaluation included standard metrics in the model decoding, and decoding, and distribution noments, with similar inference. Three dadgorithms algorithms algorit		-					
Components Components Components Components Components Components Components Code		-		•	1		•
generates functionally clear, accurate, and comprehensive code code descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVAN ÜVEGES AND BERT-Model for ORSOLY A RING ORSOLY Classifying A RING Sentiment and et al. Emotion in (2021) Political Communication Lear, accurate, and comprehensive code code code descriptions, alleviating inference. These training and distribution threads algorithms use multi-head attention, positional encoding, multi-head attention, feed-forward networks, BertC, BertNL, and Hadamard & Norm, as well as feed- forward neutral networks BERT-Model for ORSOLY Classifying A RING Sentiment and et al. Emotion in (2021) Political Communication API. Evaluation inference. These using inference. These undistribution tirends to the Java dataset. Java da				-	-		· · · · · · · · · · · · · · · · · · ·
functionally clear, accurate, and comprehensive code code descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVÁN ÜVEGES AND BERT-Model for ORSOLY Classifying A RING Sentiment and et al. (2021) Political communication (2021) Political communication (2021) Political communication in (2021) Political communication in cluded standard metrics and comprehensive code code code code (2021) Political clear of the comprehensive code (2021) Political clear of the comprehensive code (2021) Political clear of the comprehensive code (2021) Political clear of the components: the components: the components: the components: the components: the algorithms use multi-head attention, positional encoding, multi-head attention, feed-forward networks, BertC, BertNL, and Hadamard & Norm, as well as feed-forward neural networks BERT-Model for or sentiment and emotion classification, with parameters adjusted using the transformers library's Trainer API, Evaluation included standard metrics and proposed the transformers library's Trainer API, Evaluation included standard metrics and proposed the fine code distribution trends to the algorithms use multi-head attention, positional encoding, multi-head attentio		(=0=0)	*				_
clear, accurate, and comprehensive code code code descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVÁN ÜVEGES AND BERT-Model for ORSOLY A RING et al. Emotion in (2021) Political Emotion in (2021) Political Emotion in (2021) Political Emotion in Classification, with parameters library's Trainer API, Evaluation in included standard metrics Clear, accurate, and brunctional comprehensive and eligorithms use multi-head attention, positional encoding, multi-head attention, positional encoding, multi-head attention, feed-forward networks, BertC, BertNL, and Hadamard & Norm, as well as feed-forward neural networks BertNL, and Hadamard & Norm, as well as feed-forward neural networks Bert Model, Fine-Tuned for sentiment and emotion classification, with parameters adjusted using the transformers library's Trainer API, Evaluation included standard metrics Evaluation of the fine transformers Eval			~		_		
and comprehensive code code code code descriptions, alleviating issues like code understanding and achieving remarkable performance. IZ ISTVÁN performance: ISTVÁN			•				
comprehensive code code code code code code code cod			· ·	•			
code descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVÁN ÜVEGES AND BERT-Model for ORSOLY A RING ORSOLY Classifying et al. (2021) Political (2021) Political Communication Communication Code, and Decoder. Encoder, and Decoder. Decoder. Encoder, and Decoder. Decode					-	sava dataset.	- ·
descriptions, alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVÁN HunEmBERT: A ÜVEGES AND GRSOLY Classifying A RING ORSOLY Classifying et al. Emotion in (2021) Political Communication 13 Emotion in (2021) Political Communication 14 Communication 15 Decoder. attention, positional encoding, multi-head attention, feed-forward networks, BertC, BertNL, and Hadamard & Norm, as well as feed-forward neural networks 16 Decoder. attention, positional encoding, multi-head attention, feed-forward networks, BertC, BertNL, and Hadamard & Norm, as well as feed-forward neural networks BERT Model, Transformer- based ISEAR specific corpus, sentiment and emotion classification, with parameters adjusted using the transformers library's Trainer API, Evaluation included standard metrics			•				_
alleviating issues like code understanding and long dependency, and achieving remarkable performance. 12 ISTVAN diverges AND ORSOLY A RING et al. (2021) Political Communication (2021) Political Communication API. Evaluation included standard metrics processing and long dependency, and achieving remarkable performance. BertC, BertNL, and Hadamard & Norm, as well as feed-forward neural networks BERT Model, Transformer- Corpus, creating a domain specific corpus sentiment and emotion classification, with parameters adjusted using the transformers library's Trainer API, Evaluation included standard metrics Dostitional encoding, multi-head attention, feed-forward networks, as feed-forward networks BertC, BertNL, and Hadamard & Norm, as well as feed-forward neural networks Transformer- Corpus, creating a domain of the fine-tuning, Fine-tuning, Fine-tuning, the transformers library's Trainer API, Evaluation included standard metrics							performance.
like code understanding and long dependency, and achieving remarkable performance. BertNL, and Hadamard & Norm, as well as feed-forward neural networks BertC, BertNL, and Hadamard & Norm, as well as feed-forward neural networks AND BERT-Model for ORSOLY A RING et al. (2021) Political Emotion in classification, with parameters adjusted using the transformers library's Trainer API. Evaluation included standard metrics Communication Cassification Cas			•	Becouci.	*		
understanding and long dependency, and achieving remarkable performance. BertNL, and Hadamard & Norm, as well as feed-forward neural networks Bert O, BertNL, and Hadamard & Norm, as well as feed-forward neural networks BERT Model, Transformer-Datased Sentiment and emotion et al. (2021) Political Emotion in classification, with parameters adjusted using the transformers library's Trainer API. Evaluation included standard metrics Indicate the performance attention, feed-forward networks, Bert N., and Hadamard & Norm, as well as feed-forward neural networks BERT Model, HunEmPoli Corpus, creating a domain specific corpus (HunEmPoli) for sentiment and emotion classification, with parameters adjusted using the transformers library's Trainer API, Evaluation included standard metrics Indicate the performance. Bert N., and HunEmPoli The study involved corpus, creating a domain specific corpus, creating a domain specific corpus, creating a domain specific corpus,			_		*		
and long dependency, and achieving remarkable performance. 12 ISTVÁN ÜVEGES AND BERT-Model for ORSOLY A RING et al. (2021) (2021) Political Communication And long dependency, and achieving remarkable performance. Bert Model, Hadamard & Norm, as well as feed-forward neural networks HunEmBERT: A model was fine-tuned for sentiment and emotion classification, with parameters library's Trainer API, Evaluation included standard metrics Evaluation included standard metrics Evaluation of the fine transformers library's Trainer API. Evaluation included standard metrics Evaluation of the fine transformers Evaluation of the fine transformers library's Trainer API. Evaluation included standard metrics Evaluation of the fine transformers library's Trainer API. Evaluation included standard metrics					٠.		
dependency, and achieving remarkable performance. BertNL, and Hadamard & Norm, as well as feed-forward networks I2 ISTVÁN HunEmBERT: A ÜVEGES Fine-Tuned AND BERT-Model for ORSOLY Classifying A RING Sentiment and et al. Emotion in (2021) Political Communication Communication Communication Communication Defed-forward networks Bert Model, Transformer- based ISEAR specific corpus (HunEmPoli) for sentiment and emotion sentiment and emotion sentiment and emotion music included standard metrics) Dataset (HunEmPoli) for sentiment and emotion analysis in Hungarian political texts, follower by fine-tuning the huBERT model for sentiment and emotion classification task: using this corpus Evaluation of the fine			_				
achieving remarkable performance. BertNL, and Hadamard & Norm, as well as feed-forward neural networks I2 ISTVÁN HunEmBERT: A ÜVEGES Fine-Tuned AND BERT-Model for ORSOLY Classifying A RING Sentiment and et al. Emotion in (2021) Political Communication Emotion in (2021) Political Communication Emotion in (2021) Political Communication API. Evaluation included standard metrics Intervorks, BertC, BertNL, and Hadamard & Norm, as well as feed-forward neural networks BERT Model, Transformer-Lorpus, Corpus, Creating a domain (ISEAR specific corpus sentiment and emotion sentiment and emotion danalysis in Hungarian analysis in Hungarian political texts, followed by fine-tuning the hubert included standard metrics Evaluation Metrics BERT Model, HunEmPoli The study involved creating a domain (ISEAR specific corpus sentiment and emotion danalysis in Hungarian analysis in Hungarian political texts, followed by fine-tuning the hubert included standard metrics Evaluation of the fine			_		*		
remarkable performance. BertC, BertNL, and Hadamard & Norm, as well as feed-forward neural networks 12 ISTVÁN HunEmBERT: A The huBERT BERT Model, Transformer- based ISEAR specific corpus, Corpus, Creating a domain specific corpus and product tuned for ORSOLY Classifying sentiment and emotion et al. Emotion in classification, (2021) Political Communication Political with parameters Communication Political with parameters library's Trainer API, Evaluation included standard metrics BertNL, and Hadamard & Norm, as well as feed-forward neural networks BERT Model, Transformer- Corpus, creating a domain ISEAR specific corpus, ISEAR specific corpus, Sentiment and emotion classification, Trainer API, Evaluation by fine-tuning the transformers library's Trainer API, Evaluation included standard metrics							
performance. BertNL, and Hadamard & Norm, as well as feed-forward neural networks I2 ISTVAN HunEmBERT: A The huBERT model, Transformer-based pipeline, Political Communication Emotion in Classification, with parameters adjusted using the transformers library's Trainer API, Evaluation included standard metrics BertNL, and Hadamard & Norm, as well as feed-forward neural networks BERT Model, Transformer-based pipeline, Pipeline, Fine-tuning, Trainer API, Evaluation analysis in Hungarian political texts, followed by fine-tuning the huBERT model for sentiment and emotion classification task using this corpus Evaluation of the fine			~		-		
Hadamard & Norm, as well as feed-forward neural networks 12 ISTVÁN HunEmBERT: A ÜVEGES Fine-Tuned AND BERT-Model for ORSOLY Classifying A RING et al. Emotion in (2021) Political Communication Communication Hadamard & Norm, as well as feed-forward neural networks BERT Model, Transformer-based ISEAR specific corpus (HunEmPoli) for sentiment and emotion classification, with parameters adjusted using the transformers library's Trainer API, Evaluation included standard metrics					*		
Norm, as well as feed- forward neural networks 12 ISTVÁN HunEmBERT: A The huBERT model, was fine- AND BERT-Model for ORSOLY Classifying sentiment and emotion et al. (2021) Political Communication Communication Communication API. Evaluation included standard metrics Norm, as well as feed- forward neural networks BERT Model, HunEmPoli The study involved creating a domain specific corpus. ISEAR specific corpus creating a domain based ISEAR specific corpus. ISEAR specific corpus sentiment and emotion fine-tuning, analysis in Hungarian political texts, followed by fine-tuning the hubert model for sentiment and emotion classification task using this corpus Evaluation of the fine			perrormanee.				
as feed- forward neural networks 12 ISTVÁN ÜVEGES AND BERT-Model for ORSOLY A RING et al. (2021) Political Communication API. Evaluation as feed- forward neural networks BERT Model, Transformer- based Pipeline, Fine-tuning, Trainer API, Evaluation Metrics BERT Model, Transformer- based ISEAR Specific corpus (HunEmPoli) fo sentiment and emotion analysis in Hungarian political texts, follower by fine-tuning the huBERT model fo sentiment and emotion classification Metrics API. Evaluation included standard metrics							
forward neural networks 12 ISTVÁN HunEmBERT: A ÜVEGES Fine-Tuned AND BERT-Model for ORSOLY Classifying A RING et al. Emotion in (2021) Political Communication And the communication Com					*		
12 ISTVÁN							
12 ISTVÁN HunEmBERT: A The huBERT BERT Model, UVEGES Fine-Tuned model was fine-AND BERT-Model for ORSOLY Classifying sentiment and emotion et al. Emotion in Classification, (2021) Political Communication Communication API. Evaluation included standard metrics Pine-tuning the sentiment and emotion classification at a computation bears of the transformers and provided involved to the bubble and the bubble and the study involved creating a domain specific corpus, and domain sp							
12 ISTVÁN HunEmBERT: A The huBERT model, Transformer-based pripeline, Corpus, Corpus, Sentiment and emotion et al. (2021) Political Communication Communication Trainer API, Evaluation included standard metrics Pine-tuning adjusted using the transformers library's Trainer API. Evaluation included standard metrics Pine-tuning and emotion classification tasks using this corpus tasks and ard metrics Pine-tuning and pripeline, Dataset (HunEmPoli) for sentiment and emotion analysis in Hungarian political texts, follower by fine-tuning the hubERT model for sentiment and emotion classification tasks using this corpus Evaluation of the fine							
ÜVEGES AND ORSOLY A RING et al.Fine-Tuned BERT-Model for CRSOLY A RING et al.model was fine- tuned sentiment and classification, with parameters library's Trainer API. Evaluation included standard metricsTransformer- based Pipeline, Fine-tuning, Trainer API, Evaluation MetricsCorpus, specific Datasetcreating a domain specific (HunEmPoli) sentiment and emotion analysis in Hungarian political texts, followed by fine-tuning the huBERT model fo sentiment and emotion classification texaluation of the fine	12	ISTVÁN	HunEmBERT: A	The huBERT		HunEmPoli	The study involved
AND ORSOLY Classifying sentiment and emotion et al. (2021) Political Communication Communication API. Evaluation included standard metrics Pipeline, ORSOLY Classifying Sentiment and emotion classification, With parameters adjusted using the transformers library's Trainer API. Evaluation included standard metrics ISEAR Dataset (HunEmPoli) for sentiment and emotion analysis in Hungarian political texts, followed by fine-tuning the classification task using this corpus Evaluation of the fine	1-						,
ORSOLY Classifying sentiment and emotion et al. (2021) Political Communication Communication API. Evaluation included standard metrics Sentiment and emotion analysis in Hungarian political texts, follower by fine-tuning the hubers sentiment and emotion classification task using this corpus Evaluation of the fine							
A RING et al. (2021) Emotion in Political Communication Communication Emotion in Political Communication Evaluation API. Evaluation included standard metrics Evaluation of the fine Emotion in Classification, Trainer API, Evaluation included standard metrics Evaluation Fine-tuning, Trainer API, Evaluation included standard metrics Evaluation included Emotion classification, Trainer API, Evaluation included standard metrics Evaluation of the fine							*
et al. (2021) Political with parameters API, Evaluation Metrics Communication API. Evaluation included standard metrics Evaluation Trainer API, Evaluation Metrics API. Evaluation included standard metrics Evaluation Trainer API, Evaluation by fine-tuning the huBERT model for sentiment and emotion classification tasks using this corpus Evaluation of the fine			, ,		* '		` /
Political with parameters adjusted using the transformers library's Trainer API. Evaluation included standard metrics With parameters adjusted using the transformers library's Trainer API. Evaluation included standard metrics Evaluation Metrics Dolitical texts, followed by fine-tuning the huBERT model for sentiment and emotion classification task using this corpus Evaluation of the fine							
Communication adjusted using the transformers library's Trainer API. Evaluation included standard metrics Metrics by fine-tuning the huBERT model fo sentiment and emotion classification task using this corpus Evaluation of the fine					· ·		-
the transformers library's Trainer API. Evaluation included standard metrics huBERT model fo sentiment and emotion classification tasks using this corpus Evaluation of the fine		(2021)		•			· ·
library's Trainer API. Evaluation included standard metrics sentiment and emotion classification task using this corpus Evaluation of the fine			2 Jiiiii Guilloii		1.104105		-
API. Evaluation included using this corpus standard metrics classification task using this corpus Evaluation of the fine							
included standard metrics using this corpus Evaluation of the fine				-			
standard metrics Evaluation of the fine							
							1
like Precision tuned models				like Precision			tuned models'
and Recall,							
ISSN 1881 0420				and recuir,			, ////



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

	1					
13	Heewon	Reinforced	alongside error analysis using confusion matrices and manual inspection of misclassified sentences The model uses a	The model	Gigaword	demonstrated state-of- the-art results in sentiment analysis and acceptable results in emotion analysis within the political domain, accompanied by error analysis to identify common error patterns and challenges. The study presents two
	Jang and Wooju Kim et al. (2021)	Abstractive Text Summarization with Semantic Added Reward	sequence-to-sequence attention mechanism with a single-layer bi-directional LSTM encoder and decoder, incorporating intra-decoder attention to mitigate repeated phrase generation and copying difficult words.	employs policy learning and reinforcement learning to optimize metrics, focusing on word similarity, semantic similarity, and word mover distance, ensuring readability.	summarizatio n dataset.	reward functions, ROUGE-SIM and ROUGE-WMD, which add semantic values to n-gram matching. The models perform better than sequence-to- sequence, Transformer- based pre-learning, and reinforcement learning models, with improvements in readability and grammatical accuracy
14	Yan Du and Hua Huo et al. (2020)	News Text Summarization Based on Multi- Feature and Fuzzy Logic	The proposed method uses a multi-step approach for automatic news text summarization, including preprocessing, segmentation, stop-word removal, tokenization, stem extraction, and extraction of key features at word and sentence levels.	method for extracting word features using a SpeGenetic Algorithm, a fuzzy logic system, and a specific keyword extraction method.	DUC2002 Datasets.	This paper proposes a new model combining multi-feature, genetic algorithm, and fuzzy logic for news text summarization. It extracts important features like word and sentence features, graded based on news text characteristics, and uses a fuzzy logic system to generate fair weights. The method uses ROUGE-1 for fitness values and ROUGE-2 for optimal weight allocation. Experimental results



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

						outperforms other
						methods in generating
						high-quality news
						summaries.
15	Rini	CE-BERT:	The	Uses various	Twitter 15	The study reveals CE-
	Anggraini	Concise and	methodology of	layer	Dataset,	BERT is an efficient
	ngsih;	Efficient BERT-	the study	reduction	Twitter 16	Twitter rumor detection
	Ghulam	Based Model for	involves	strategies,	Dataset and	model with reduced
	Mubashar	Detecting	proposing a	including the	PHEME	computational
	Hassan;	Rumors on	concise and	basic model,	Dataset	requirements,
	Amitava	Twitter	efficient BERT-	six-layer		outperforming state-of-
	Datta et.al		based model	models, four-		the-art models in source
	(2023)		named CE-	layer models,		text scenarios.
			BERT for	and two-layer		
			detecting rumors	models, and		
			on Twitter. The	emphasizes		
			key techniques	the		
			employed in	importance of		
			developing CE-	a fine-tuning		
			BERT include	process.		
			selecting			
			effective layers,			
			efficient fine-			
			tuning, and			
			classification			
16	YOUHYU	Multi-Encoder	Transformer-	Transformer	Law (AI-	The paper introduces a
	N SHIN et	Transformer for	Based Encoder-	architecture,	Hub) dataset	transformer-based
	al.	Korean	Decoder Model,	multi-encoder	News (AI-	encoder-decoder model
	(2023)	Abstractive Text	Multi-Encoder	architecture,	Hub) dataset	tailored for Korean
		Summarization	Architecture,	fine-tuning	News (NIKL)	abstractive
			Combining	with BERT-	dataset	summarization and
			Representations,	based PLMs,		proposes a multi-
			Fine-Tuning,	Flat and		encoder architecture
			Auto-regressive	parallel		utilizing various pre-
			Decoding,	combination		trained models. It
			Evaluation	strategies for		demonstrates
			Metrics.	combining		significant performance
				representation		enhancements over
				s, Auto-		single-encoder models,
				regressive		particularly with
				decoding,		diverse encoders,
				Evaluation		across three Korean
				using		summarization datasets,
				ROUGE and		while also conducting
				BERTScore		thorough ablation
				metrics		studies and qualitative
						analyses to assess the
						efficacy of different
					187	ISSN IN NO.



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

	T					
						techniques, including OOV token handling.
17	Fucheng	A Topic	TIF-SR is a text	The model	LCSTS and	The TIF-SR model
	You,	Information	summarization	uses BERT-	NLPCC2017	significantly improves
	Shuai	Fusion and	method that	based	Datasets	ROUGE scores on
	Zhao, and	Semantic	combines topic	encoding,		LCSTS and
	Jingjing	Relevance for	information and	Topic		NLPCC2017 datasets
	Chen et.al	Text	semantic	Information		by incorporating topic
	(2020)	Summarization	relevance,	Fusion,		information and
			enhancing the	Transformer		semantic relevance,
			quality of	architecture,		resulting in fluent,
			generated	semantic		readable, and
			summaries by	relevance calculation,		semantically relevant
			minimizing negative log	and a fine-		summaries.
			probability and	tuning		
			maximizing	objective to		
			cosine similarity.	enhance		
			cosme similarity.	semantic		
				relevance in		
				source		
				documents.		
18	YUSUNG	Topic Modeling-	The	The	General	The paper achieved
	AN,	Based	methodology	techniques	datasets are	significant
	DONGJU	Framework for	used in the paper	employed	considered	advancements in
	KIM,	Extracting	involves	include	for survey	extracting product
	JUYEON	Marketing	extracting	clustering	paper.	characteristics,
	LEE,	Information from	strengths and	algorithms,		identifying related
	HAYOUN	E-Commerce	weaknesses of	topic		products, and
	G	Reviews	individual	modeling		forecasting product
	OH,JOO-		products,	(such as		trends using topic
	SIK LEE,		identifying	BERTopic		modeling techniques
	DONGH		connections	and DCN),		and a Transformer-
	WA		among different	and a		based model.
	JEONG et		products, and	Transformer- based		Moreover, it provided
	al. (2023)		predicting product trends.	forecasting		valuable insights into consumer preferences
	(2023)		product trends.	model.		and market trends
				It		through thorough
				demonstrated		analysis and evaluation
				the		
				effectiveness		
				of employing		
				multiple		
				techniques		
				such as		
				BERTopic,		
				DCN, and		BEARCH IN SCHOOL
					1187	1211



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

			- ,	oue i, i ebildaiy	T	
				Transformer		
				models,		
				showcasing		
				improvement		
				s over		
				traditional		
				methods like		
				LDA and K-		
				means		
				algorithms		
10	т	A DEDT D 1	TTI 1		TP1 1 4	TI 1 MID
19	Lavanya	A BERT Based	The research	Techniques	The dataset	The research uses NLP,
	Elluri, Sai	Approach to	proposes a	like Entity	comprises	deep learning, and
	Sree Laya	Measure Web	framework for	Recognition,	over 3000	semantic web
	Chukkapal	Services Policies	automatic	Modal Logic,	web service	techniques to
	li, Karuna	Compliance	context	Machine	provider	efficiently analyze
	Pande	With GDPR	extraction and	Learning,	privacy	short or incomplete
	Joshi, Tim		comparison of	BERT Text	policies,	texts, focusing on
	Finin, and		short text	Summarizatio	categorized	privacy policies in
	Anupam		documents,	n, and	into EU-	progress.
	Joshi et al.		specifically	Semantic	based and	F8
	(2021)		focusing on	Web	worldwide	
	(2021)		service provider	Languages	policies,	
					*	
			policies,	are used to	downloaded	
			specifically	extract high-	post-May	
			applying GDPR	frequency	2018 to	
			and comparing	entities from	ensure	
			with 3000 web	GDPR	relevance	
			service privacy	documents.	post-GDPR.	
			policies. The			
			framework			
			includes six			
			phases.			
20	Shuai	Leveraging Pre-	The research	The study	LCSTS	The proposed BSA
	Zhao,	Trained	proposes a	explores	dataset	model outperforms
	Fucheng	Language Model	Chinese text	pretrained	anuse	baseline models on the
	You, and	for Summary	summary model	language		LCSTS dataset,
	-	•	•			
	Zeng	Generation on		models,		<u> </u>
	Yuan et	Short Text	templates, fine-	extractive		scores and text division
	al.		tuning with	models, and		quality. Future work
	(2020)		pretrained	abstractive		should explore its
			language models	models for		applicability to other
			like BERT,	language		pretrained language
			involving data	understanding		models.
			preprocessing,	tasks,		
			model	extractive		
			architecture, and	summarizatio		
			experimentation.	n, and		
			p	abstractive		RSEARCH IN SCIEN
				abbuuchve	l //oS	



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

		1		oue i, i obradiy		
				summarizatio		
				n,		
				incorporating		
				attention		
				mechanisms		
				and BERT.		
21	Sung-Guk	Learning Cluster	The text	The study	CNN/DailyM	A new pre-trained
	Jo, Seung-	Patterns for	discusses a	uses various	ail Dataset	sequence-to-sequence
	Hyeok	Abstractive	proposed	techniques	and XSUM	model improves
	Park,	Summarization	abstractive	for	Dataset	abstractive
	Jeong-Jae		summarization	abstractive		summarization by
	Kim, and		model that builds	summarizatio		extracting salient
	Byung-		upon the BART	n, including		sentences and
	Won et al.		architecture by	the Encoder-		clustering context
	(2016)		adding a	Decoder		vectors. Experimental
			discriminator	Attention		results outperform
			and a cluster	Model,		existing models,
			generator. The	Pointer		addressing fact
			model splits	Network,		inconsistency and long
			original text into	Coverage		document
			sentences and	Mechanism,		summarization.
			classifies each	Transformer		Sammarization.
			sentence as	Model,		
			salient or non-	UniLM,		
			salient, creating	BERTSUM		
			clusters for	and BART		
			salient and non-	Models,		
			salient sentences.	PEGASUS		
			Experiments on	Model,		
			XSUM and	Reinforcemen		
			CNN/DailyMail	t Learning,		
			datasets showed	Topic Model,		
			the model	Multimodal		
			outperformed	Information,		
			BART and	Attention		
			PEGASUS in	Head		
			terms of ROUGE	Masking,		
			and BERTScore.	Information		
				Theory,		
				Extraction-		
				and-		
				Paraphrasing,		
				Entity		
				Aggregation		
				and		
				Factuality		
				Consistency,		
				Deep	le.	SEARCH III SCIENCE
				Communicati		ISSN
Conv	right to IJAR	SCT	DOI: 10.48175/IJ	ARSCT-15356	1/2	581-9429 (8) 425

Copyright to IJARSCT www.ijarsct.co.in



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

				ng Agents, Sentence Corresponden ce, Graph, and Bottom- Up Approach.		
22	Bihui Yu; Chen	Policy Text Classification	The algorithm first employs a pre-trained	The technique used in the	The policy text data, documents	The paper achieves an improvement in policy
	Deng; Liping Bu et al.	Algorithm Based on Bert	BERT language model to	paper is BERT	from various policy	text classification accuracy, with the trained model
	(2022)		vectorize the sentence-level	(Bidirectional Encoder	domains	achieving the highest F1 value of 93.25% on
			features of the policy text.	Representations from		the test set. This result demonstrates a nearly
			Then, the	Transformers		6% increase in
			obtained feature vectors are input), a transformer-		performance compared to the BERT model's
			into a classifier	based		classification task for
			for classification	language model pre-		the MRPC (Microsoft Research Paraphrase
				trained on		Corpus) task.
				large corpora of text data.		
23	Songsong	Text	The	The	Text data	The paper achieves an
	Liu;	Classification	methodology	techniques	related to	improvement in the
	Haijun	Research Based	used in the paper	used in the	people's	accuracy of text
	Tao;	on Bert Model	involves	paper are the	livelihood	classification by
	Shiling	and Bayesian	combining the	Bert model, a	governance,	combining the
	Feng et al.	Network	Bert model with	pre-training	containing	Bayesian network's
	(2019)		Bayesian	model based	unstructured	coarse classification
			networks for text classification.	on deep learning for	data from various	with the Bert model's
			Initially, the	natural	sources	fine-grained classification. This
			Bayesian	language	within this	combination reduces
			network is	processing	domain.	errors caused by
			utilized to	(NLP), and		classification defects
			perform a coarse	Bayesian		inherent in using only
			classification	networks, a		one of the methods,
			into two	probabilistic		ultimately enhancing
			categories,	graphical		the overall accuracy of
			providing an approximate	model that represents a		text classification.
		1	approximate	_		
			category range	set of random		l
			category range for each text.	set of random variables and		
			category range for each text. Subsequently,	set of random variables and their		
			for each text.	variables and		
			for each text. Subsequently,	variables and their		TO ACCO IN EAST



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53 Volume 4, Issue 1, February 2024

			classify the text into specific categories within	via a directed acyclic graph.		
24	Younghoo n Seo, Jihyeok Park, Gyungtae k Oh, Hyungjoo Kim, Jia Hu, and Jaehyun So et al. (2023)	Text Classification Modeling Approach on Imbalanced- Unstructured Traffic Accident Descriptions Data	the identified range. The methodology classifies and interprets unstructured traffic crash description data through three stages: data preprocessing, model development, and model performance evaluation. It uses wordembedding algorithms, a BERT-based text classification	Word Embedding Algorithm, NLP Models, BERT-Based Text Classification Model, K- Fold Cross- Validation, and Stratified K-Fold Cross- Validation for text preprocessing , data quality enhancement, and classification.	The study analyzes 2,427 traffic crash datasets from the Korean Traffic Crash Analysis System in Daegu, Korea, focusing on crash data elements, weather conditions, and written descriptions.	The study developed a method for classifying traffic crash text into standardized data using the BERT model, a natural language processing technique. The model exhibited the highest performance in interpreting three traffic crash elements and had a text classification accuracy of over 95% for most labels. The study aims to increase the use of underutilized traffic crash data in traffic safety research, reducing crashes and
			model, and cross-validation for robustness.			improving traffic safety.
25	Xuedong Tian and Jiameng Wang et al. (2021)	Retrieval of Scientific Documents Based on HFS and BERT	The scientific document retrieval system consists of three modules: Document Preprocess, Mathematical Expression Similarity, and Text Similarity, which extract mathematical expressions, unify LaTeX and MathML, and calculate similarity between text parts.	Techniques used include FDS for decompositio n, HFS for calculating similarity, BERT for text similarity, keyword extraction, and cosine similarity to analyze mathematical expressions.	The dataset consists of scientific documents, containing mathematical expressions and related texts. Keywords from these contexts were used to train BERT and calculate text similarity.	This study proposes a retrieval method combining mathematical expressions and contextual text for mathematical information retrieval. It calculates membership degree, HFS, BERT, context keywords, and word vectors. Future studies aim to improve extraction methods, address special cases, and evaluate Chinese and English algorithms.



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 1, February 2024

An, Based automates utilized analyzed modeling Dongju Framework for product clustering 530,877 product c	used topic to extract haracteristics ed products
Dongju Framework for product clustering 530,877 product c extraction, pros algorithms, consumer and related to the form of the consumer and constant	haracteristics
Kim, Extracting extraction, pros algorithms, consumer and related and cons, and dimension reviews from from	
Juyeon Marketing and cons, and dimension reviews from from	r Production
	consumer
	aiming to
Hayoung E-Commerce from Naver algorithms, Shopping for uncover	marketing
Oh, Joo- Reviews Shopping review topic 17 products, insights.	The DCN-
	el performed
	its results
	ghtly less
Jeong are Transformer- n, and analyzed favorable	than the
all based forecasting transformer- search baseline	BERTopic.
	ts can guide
of the Mecab, POS forecasting to trends using marketing	strategies,
	oducts with
	entions, and
(2023) Modeling, and extract clustering inform	product
co-occurrence insights, and algorithms. introduction	•
frequency to predict product	trend
	model also
	ts predictive
products.	-
	key product
	from text
	owever, it
	· ·
changes	edict future due to
identifying	
	past data
27 Multi-Label The Preprocessing MLBioMedL The paper	
	presents a
	nodology for
	biomedical
	nto multiple
	esulting in
MAHMO in Biomedical components: dependency enhanced	accuracy
	sophisticated
MUHAM Answering Biomedical Feature preprocess	
MAD Question extraction, classificati	
NABEEL Classification Data techniques	
ASIM, and Proposed Transformati Furthermon	
MUHAM Biomedical on, proposed	question
MAD Question Classification answering	approach
USMAN Answering , Collective demonstrate Methodology Pageoling Taggreen	
	e compared
et. Al to existing to existing to existing the existing to exist the existing the existing the existing the existing the exist	
(2019) surpassing	baseline
	showcasing
Copyright to IJARSCT DOI: 10.48175/IJARSCT-15356	icacy in 428

Copyright to IJARSCT www.ijarsct.co.in

DOI: 10.48175/IJARSCT-15356

428



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

					I	biomedical question
						*
20	0: 1:	Electric Power	TI 4 1 1	EDAT DEDT	TPI 4 1	answering tasks.
28	Qinglin		The study aims	EPAT-BERT	The study	The EPAT-BERT
	Meng,	Audit Text	to efficiently	introduces	uses a dataset	model, based on the
	Yan Song,	Classification	classify electric	two pre-	of 1.5M	classical pre-training
	Jian Mu,	With Multi-	power audit texts	training tasks	electric	language model BERT,
	Yuanxu	Grained Pre-	using pre-	for electric	power-related	focuses on two pre-
	Lv,	Trained	training and fine-	power audit	texts, 1,500	training tasks: word-
	Jiachen	Language Model	tuning models	text	audit texts,	granularity masked
	Yang,		like BERT,	classification	and baseline	language model and
	Liang Xu,		MAE, and CLIP,	using BERT's	models like	entity-granularity
	Jin Zhao,		which require	structure. The	Naive Bayes,	masked language
	Junwei		less supervised	tasks include	SVM, GBDT,	model. The model's
	Ma, Wei		data for fine-	word-	AdaBoost,	effectiveness in
	Yao, Rui		tuning and	granularity	XGBoost,	classifying electric
	Wang, and		outperform non-	masked	TextCNN,	power audit texts is
	Maoxiang		pre-trained	language	LSTM, and	demonstrated through
	Xiao		models.	modeling and	BERT.	two ablation
	et .al			entity-		experiments. The
	(2023			granularity		results show that the
				masked		model significantly
				language		exceeds existing
				modeling,		models in accuracy,
				enhancing		precision, recall, and F1
				understanding		score compared to
				of electric		traditional machine
				power audit		learning models. The
				texts.		model's extensibility
						allows it to be applied
						to other electric power
						text classification tasks
						and project type
						annotation.
29	Wafaa	Automatic Text	The	The paper	Survey paper	The paper achieves the
	Samy El-	Summarization:	methodology	covers	does not	goal of providing
	Kassas,	A	employed in this	extractive,	contain any	researchers with a
	Cherif	Comprehensive	research involves	abstractive,	datasets.	comprehensive
	Salama,	Survey	conducting a	and hybrid		overview of ATS,
	Ahmed		comprehensive	approaches to		serving as a valuable
	Rafea,		survey of	ATS,		resource for
	Hoda K.		Automatic Text	highlighting		understanding the
	Mohamed		Summarization	their		current state of the
	et.al		(ATS)	differences		field, identifying
	(2019)		techniques and	and		research gaps, and
			methodologies.	applications.		guiding future
			The paper	Additionally,		directions for
			discusses various	it explores		improving ATS
			approaches,	different		technique
_	· · · · · · · · · · · · · · · · · · ·	·	·	·	1/3/	ISSN /III

Copyright to IJARSCT www.ijarsct.co.in



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

	methods,	building
	building blocks,	blocks and
	techniques,	techniques
	datasets,	used in ATS,
	evaluation	along with
	methods, and	the datasets
	future research	commonly
	directions in	employed for
	ATS.	training and
		evaluation
		purposes.

III. CONCLUSION

According to the literature survey on tracking devices for Alzheimer's patients, there is a growing interest in utilizing technology to meet the special issues related with the care and safety of those suffering from Alzheimer's disease. The studies constantly highlight the ability of monitoring devices to improve the overall quality of life for both patients and carers by providing answers to issues such as moving around, safety concerns, and timely reaction during emergencies. Several study studies have looked into the usefulness of various tracking technologies, such as GPS-based systems, wearable devices, and IoT-enabled solutions. These technologies not only help with location tracking but also include features like fall detection, geofencing, and conduct monitoring. The literature stresses the necessity of user-friendly interfaces and customizing to meet the special demands of Alzheimer's patients, ensuring that the devices are not intrusive and acceptable to the users. Interdisciplinary cooperation among healthcare experts, technologists, and carers are critical in the development and implementation of tracking devices. The integration of advanced sensors, data analytics, and communication technologies enables a holistic approach to addressing Alzheimer's disease difficulties, including real-time information and quick responses. Finally, the literature review stresses the potential of tracking devices to dramatically improve the care and safety of Alzheimer's patients. Future research and development efforts should focus on addressing existing challenges in order to create improved, user-friendly, and ethically sound tracking solutions that improve the lives of Alzheimer's patients while also reducing the burden on their caretakers..

REFERENCES

[1] Ayesha Ayub Syed, Ford Lumban Gaol, Tokuro Matsuo et al. (2020) "A Survey of the State-of-the-Art Models in Neural Abstractive Text Summarization"

[2]Lochan Basya, Mihir Sanghvi et al. (2022) "Text Summarization Using Large Language Models: A Comparative Study of MPT-7b-instruct, Falcon-7b-instruct, and OpenAI Chat-GPT Models"

[3]Pancheng Wang, Shasha Li, Kunyuan Pang, Liangliang He, Dong Li, Jintao Tang, Ting Wang et al. (2023) "Multi-Document Scientific Summarization from a Knowledge Graph-Centric View"

[4]Shuai Zhao, Fucheng You, and Zeng Yuan Liu et al. (2022) "Leveraging Pre-Trained Language Model for Summary Generation on Short Text"

[5]Ming Zhong, Yang Liu, Suyu Ge, Yuning Mao, Yizhu Jiao, Xingxing Zhang, Yichong Xu, Chenguang Zhu, Michael Zeng, Jiawei Han et al. (2022) "Unsupervised Multi-Granularity Summarization"

[6]Zhixin Li, Zhi Peng, Suqin Tang, Canlong Zhang, Huifang Ma et al. (2020) "Text Summarization Method Based on Double Attention Pointer Network"

[7]Pratik K. Biswas 1 and Aleksandr Iakubovich et al. (2022) "Extractive Summarization of Call Transcripts"

[8]Ángel Hernández-Castañeda, René Arnulfo García-Hernández2 and Yulia Ledeneva et al. (2023) "Toward the Automatic Generation of an Objective Function for Extractive Text Summarization"

[9]Fucheng You, Shuai Zhao, and Jingjing Chen et al. (2020) "A Topic Information Fusion and Semantic Relevance for Text Summarization"

[10]Luca Cagliero, Laura Farinetti, and Elena Baralis et al. (2019) "Recommending Personalized Summaries of Teaching Materials"

[SSN | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 2581-9429 | 25

DOI: 10.48175/IJARSCT-15356

Copyright to IJARSCT www.ijarsct.co.in

430

IJARSCT



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 4, Issue 1, February 2024

- [11]Ruyun Wang, Hanwen Zhang, Guoliang Lu, Lei Lyu, and Chen Lyu et al. (2020) "Fret: A Reinforcer-Transformer Model for Code Summarization"
- [12]István Üveges and Orsolya Ring et al. (2021) "HunEmBERT: A Fine-Tuned BERT-Model for Classifying Sentiment and Emotion in Political Communication"
- [13]Heewon Jang and Wooju Kim et al. (2021) "Reinforced Abstractive Text Summarization with Semantic Added Reward"
- [14] Yan Du and Hua Huo et al. (2020) "News Text Summarization Based on Multi-Feature and Fuzzy Logic"
- [15]Rini Anggrainingsih, Ghulam Mubashar Hassan, Amitava Datta et al. (2023) "CE-BERT: Concise and Efficient BERT-Based Model for Detecting Rumors on Twitter"
- [16] Youhyun Shin et al. (2023) "Multi-Encoder Transformer for Korean Abstractive Text Summarization"
- [17] Fucheng You, Shuai Zhao, and Jingjing Chen et al. (2020) "A Topic Information Fusion and Semantic Relevance for Text Summarization"
- [18] Yusung An, Dongju Kim, Juyeon Lee, Hayoung Oh, Joo-Sik Lee, and Donghwa Jeong et al. (2023) "Topic Modeling-Based Framework for Extracting Marketing Information from E-Commerce Reviews"
- [19] Lavanya Elluri, Sai Sree Laya Chukkapalli, Karuna Pande Joshi, Tim Finin, and Anupam Joshi et al. (2021) "A BERT Based Approach to Measure Web Services Policies Compliance With GDPR"
- [20] Shuai Zhao, Fucheng You, and Zeng Yuan et al. (2020) "Leveraging Pre-Trained Language Model for Summary Generation on Short Text"
- [21] Sung-Guk Jo, Seung-Hyeok Park, Jeong-Jae Kim, and Byung-Won et al. (2016) "Learning Cluster Patterns for Abstractive Summarization"
- [22] Bihui Yu; Chen Deng; Liping Bu et al. (2022) "Policy Text Classification Algorithm Based on Bert"
- [23] Songsong Liu; Haijun Tao; Shiling Feng et al. (2019) "Text Classification Research Based on Bert Model and Bayesian Network"
- [24] Younghoon Seo, Jihyeok Park, Gyungtaek Oh, Hyungjoo Kim, Jia Hu, and Jaehyun So et al. (2023) "Text Classification Modeling Approach on Imbalanced-Unstructured Traffic Accident Descriptions Data"
- [25] Xuedong Tian and Jiameng Wang et al. (2021) "Retrieval of Scientific Documents Based on HFS and BERT"
- [26] Yusung An, Dongju Kim, Juyeon Lee, Hayoung Oh, Joo-Sik Lee, and Donghwa Jeong et al. (2023) "Topic Modeling-Based Framework for Extracting Marketing Information from E-Commerce Reviews"
- [27] Muhammad Wasim, Waqar Mahmood, Muhammad Nabeel Asim, Muhammad Usman Ghani et al. (2019) "Multi-Label Question Classification for Factoid and List Type Questions in Biomedical Question Answering"
- [28] Qinglin Meng, Yan Song, Jian Mu, Yuanxu Lv, Jiachen Yang, Liang Xu, Jin Zhao, Junwei Ma, Wei Yao, Rui Wang, and Maoxiang Xiao et al. (2023) "Electric Power Audit Text Classification With Multi-Grained Pre-Trained Language Model"
- [29] Wafaa Samy El-Kassas, Cherif Salama, Ahmed Rafea, Hoda K. Mohamed et al. (2019) "Automatic Text Summarization: A Comprehensive Survey".

