HIGH PERFORMANCE ARTICLE WRITING USING DEEP LEARNING

[1] Dr. T.Subha Mastan Rao, [2] Paridhi Wahii, [3] Abisha Winslet Beri, [4] Suchith Adepu

[1]Associate Professor, Computer Science And Engineering, CMR Technical Campus.
[2]Student, Department of Computer Science and Engineering CMR Technical Campus.
[3]Student, Department of Computer Science and Engineering CMR Technical Campus.
[4]Student, Department of Computer Science and Engineering CMR Technical Campus.

Abstract — Content writers spend endless nights creating content which might or might not aid their requirement. A lot of times the content produced is not good in terms of qualitative nature of the entire text. Technical writing can be equally challenging The challenge of understanding topics, finding data specific to the need of topic, making sure information acquired is validated and finding time to do the research required can be cumbersome The model we propose aims to shorten the time it takes to create high-quality content. Traditional supervised state-of-the-art models, which were utilized in content generation, are based on a variety of assumptions from various research viewpoints but it produced unwanted data which lead to decelerating efficiency. Long strings of text were generated, but they lacked creativity and human sentiment, and they couldn't tell the difference between good content and crude language. The content created was often repetitive in nature. Our project aims to help in generating content for technical articles which require factual information or just reference material for writers needing good quality data in a short amount of time. We use a transformer-decoder only model also known as GPT-2, which is an unsupervised language model. It showed that training on larger dataset and having more parameters improved the capability of language model to understand tasks and surpass the supervised language model. As a language model it produces large amounts of quality copy. Through our project content can be enhanced and produced faster, while humans can intervene to insert creativity and uniqueness. Deep learning and natural language processing are the two key features of artificial intelligence (AI) technologies that help AI tools in using prevailing information about any subject from the database and processing it accordingly.

***Keywords*— Text generation, Deep Learning, Deep neural network, Gpt-2 ,Artificial Intelligence, Transformer, Transformer-decoder , Streamlit , Scrapy, Feed-forward neural network**

I. INTRODUCTION

This project is titled as "High Performance Article Writer Using Deep Learning" which is trained on a dataset that comprises of thousands of research paper abstracts. It can write articles for business blogs, technical writing, websites, corporate communication, research papers, technical articles, etc. The Model is trained to curate content for technical writing and follow a formal tone. The proposed system takes the title (keyword or string of keywords) to identify and recommend content that writers rework or reference to produce new content. It takes existing content or assists with content development and rewrites it to create new content. It aids in the preparation of project papers, journalistic articles, or simply reference material for authors while avoiding plagiarism issues. It's takes factual information and presents in a digestible form — relevant to the audience. It aims to save time and effort. As technical article requires lots of research, it can be beneficial to people who have to write lengthy articles or require reference material for their research papers. The proposed system provides authentic and validated information to the user. It is trained to be grammatically correct and understand the tone of technical writing. It provides quality and relevant content in a short span of time. Reduces effort of writer. It produces content that is specific to the title taken in as input. The length and degree of creativity can be determined by user.

II. LITERATURE REVIEW

In [1], The paper uses two popular AI models for text generation and perform a comparative analysis. OpenAI GPT-2 and BERT models are used for prediction and generation of text .GPT-2 is trained on two corpora to generate long sentences along with articles while the BERT model is used for prediction of immediate words on the basis of given context .

In[2],The paper explores the use of pre-trained NLP models, BERT and OpenAI GPT-2 to gather coronavirus related information and literature. Text summarization is performed on the COVID-19 Open Research Dataset (CORD-19). The models are used to perform mapping from a specific keyword selected to generate summary text, resulting in an abstract summary. The goal is to bring researchers closer to fast growing publications of COVID literature and uses the ROUGE metric to evaluate the text summarization.

In [3], Generative Pretrained Transformer 2 (GPT-2), an NLP model is used to experiment with text generation to test if fake reviews generated by AI tools for academic purposes can be feasible .The 355M version of GPT-2is used and fine tuned with a corpus of review reports which is based on PeerRead dataset.

In[4],The paper observes that a natural language model can perform down stream tasks in a zero shot setting without any sort of modification and when a large language model is trained on a large and diverse dataset , the model's performance increases across many domains. It is demonstrated that the language models when experimented with a large dataset begin to learn tasks without any explicit supervision. The resulting model trained on the CoQA dataset – excels in comparison with other baseline models

In[5], The paper explores way to control attributes of the generation capabilities of different transformer based language models that have been trained on large The proposed model combines a pretrained language model with simple classifiers .The attributes to control generation are use a bag of words (BoW) which relates to a topic and a linear discriminator that is trained on top of LM latent representations to control sentiment. fine-grained control of attributes is achieved by a simple gradient-based sampling mechanism

In[6], A fine-tuned model of BLEURT is used for pseudo-response selection. The dataset on which the experimentation is done is SGD and Weather benchmarks, and it is observed that the pro020 posed self-training approach improved the tree accuracy by 46%.

In [7], Global and local addressing scheme is used to structure the table content that uses a sequence to sequence technique for the model. Local addressing scheme is responsible for determining word in the table that would be used for the description generation while global addressing scheme determines word for summary generation. The LSTM model is used for generating natural language summarization of table.

The dual attention mechanism present in the decoding phase is used for generating description from the table. Experimentation is done on the WIKIBO dataset .

In,[8],The paper showcases the use of a method the authors call GENPET which is based on pattern-exploiting training which also employs supervised learning . It was experimented on PEGASUS(Transformer encoder-decoder architecture) a spyware that yielded better results when compared to common finetuning across a set of different tasks and different training set sizes .GENPET only works for classification tasks and is observed to give good results in few shot settings. The paper explains GENPET to be a finetuning procedure created for generative language models that achieved goo data efficiency and used textual instructions and training examples.

In[9], an open-source toolkit is introduced that supports the broad set of text generation tasks transforms inputs into natural language. Texar extracts common patterns from the tasks and creates a library of reusable modules, and also allows arbitrary model architectures and algorithmic paradigms. The toolkit supports lot of large-scale pretrained models, TensorFlow and PyTorch. It is released under Apache License 2.0

In[10], The paper discusses how Pre-trained transformer models continue to increase in size and the different approaches to compress or utilize large pre-trained checkpoints into smaller and faster versions which retain the performance of the original models .Researches also show that subsets of trained teacher models can be extracted without affecting performance and therefore a "shrink and fine-tune" ("SFT") approach is discussed that extracts a student model from the maximally spaced layers of a fine-tuned teacher model . The hypothesis is that removing full layers will have minimal impact performance. Shrunken model is used to run the original fine-tuning procedure without any sort of modification .Experimentation is done on CNN and XSUM datasets.

In[11], The SLR a review paper that discusses 5 research aspects that associate with text generation. It reviews the quality metrics use for evaluating generated text, datasets used for training, languages on which the text generation is performed, and applications. Main aim of the survey is put together the relevant work in a systematic order and highlight important contributions from different researchers focusing on the past, present, and future trends. 90 primary studies are identified from 2015 to 2021 using the PRISMA framework.

In[12], Working of Bidirectional RNN is explored which makes use of two RNN layers and the sequence is looked in both forward and backward directions and output is combined .It is observed that BRNN can be trained without limiting it to using the input information. Experimentation uses the TIMIT database.

In[13],Different works on deep learning especially that associated with natural language processing are reviewed .Technologies like Recurrent Neural Networks (RNNs),Convolutional Neural Networks (CNNs),Variational Auto-Encoders (VAEs),Generative Adversarial Networks (GANs),Activation functions and Optimization techniques are explored and discussed .

In[14] a novel neural architecture is proposed which enables learning dependency without having to determine a fixed-length. It also doesn't disrupt temporal coherence It observes a dependency 80% longer than that of RNNs and 450% longer for vanilla Transformers. Proposed model shows to achieve good performance on short and long sequences, and evaluates to be faster than other baseline systems.

In[15], The paper discusses error accumulation in Neural Machine Translation. context words were sampled from ground truth sequence and predicted sequence by the model during training. The approach achieved visible improvements in multiple datasets.

In[16], The paper introduces a new language representation model called BERT, (Bidirectional Encoder Representations from Transformers) designed to pretrain deep bidirectional representations from unlabeled texts. BERT is conceptually simple and empirically powerful.

In[17], a simple approach to implement conversational modeling task is explored .It uses sequence to sequence framework. The model works by predicting the next sentence given a previous set of sentences It extracts knowledge from both a domain specific dataset, and from a large, noisy, and general domain dataset of movie subtitles. the lack of consistency is a failure mode of the proposed model.

In[18]The paper observes that large gains on tasks involved in the understanding of natural language processing can be achieved by generative pre-training of a language model on a diverse corpus of unlabeled text, which should be followed by specific fine-tuning on each task. Task-aware input transformations are used to achieve effective transfer qand causing minimal changes to the model architecture.

In[19] The paper reviews different neural approaches to conversational AI that have come up in the last few years. Three categories of conversational systems are discussed that include question answering agents, task-oriented dialogue agents, and chatbots. Techniques like DNN-based response generation, neural Machine Reading Comprehension (MRC) model, Implicit ReasoNet (IRN) and d Neural Logic Programming (Neural LP)etc are discussed .

In[20], The paper proposes a generalized autoregressive pretraining method to overcome limitations of the language model known as BERT .The proposed technique called XLNet also enables learning context in both forward and backward directions.It has autoregressive formulation. It also uses ideas from Transformer-XL for the pretraining procedure .For some cases it is observed to outperform Bert.

In[21], A two parameter-reduction technique is proposed to over the issues of speed and memory consumption with BERT. The proposed model is said to scale better than the original model of BERT and uses self-supervised loss to focuses on modeling inter-sentence coherence which helps downstream tasks with multi-sentence inputs. Results of the model on GLUE, RACE, and SQuAD benchmarks are impressive

In[22], In this work, A new task: emotion-cause pair extraction (ECPE)is proposed to extract the potential pairs of emotions and the causes of the emotion in a file.It aims to overcome the shortcomings of traditional Emotion cause extraction (ECE)models .Here ,individual emotion extraction is performed and cause extraction is done using multi-task learningafter which emotion-cause. pairing and filtering is performed .

In [23], it makes use of the pcRU framework which is language generation framework to generate weather forecast . The models were evaluated on the basis of their output quality .Improved development time, increase in reusability of systems and better computational efficiency was observed over select NLG systems and happened to reduce computational expenses . Informed decisions were made during the generation phase.

In [24], A neural model is proposed to scale large and rich domains. Purpose of the system is to create large amounts of data specific to the input image given that uses OCR to recognize each character. A web crawler is used to gather data from datasets along the web mainly validated education websites in the text generation module. Main aim is to be able to provide information comparable to that found in text books for students of different educational field by taking in the syllabus as input

In [25],Automatic text generation from given input in the form of structured data is explored.Data2Text platform mainly describes the input text in the text generated and has three components namely model training, model revision and text generation. It uses semi-HMMs model to extract high quality templates and corresponding trigger conditions from parallel data

In [26]Proposed system is an entity-centric neural architecture for data-to-text generation where the model creates entity-specific representations which are dynamically updated. Text generated as output is based on the input that goes through entity memory representations which use hierarchical attention at each step. Output generated is coherent, concise and factually correct. The ROTOWIRE benchmark dataset is used for experimentation on the baseball domain .Results of the proposed model are observed to outperform competitive baselines in automatic and human evaluation.

In [27], The paper explores the idea of how decisions for word ordering and word choice in surface natural language generation can be automatically learned from annotated data . Four trainable systems are used in which NLG1 is kept as a baseline system for comparison and treated as a lookup table while NLG2 and NLG3 attempt to find the highest probability word sequence with respect to a maximum entropy probability model.NLG4 requires a dependency -style grammar for the fragments of phrases and is kept consistent with the rules of grammar.NLG4 implements dialogue strategy where order of words are dynamically modified The proposed system reorders words at runtime is basically represents practicality of learning the decisions for ordering the word and word choice. Thereafter we see a system that makes a conversation that has a more natural tone.

In [28] ,The paper explores the idea of using an ensemble network to achieve the goalof generating conclusion-supplement answers for non-factoid questions which is a difficult for the currently existing encoder-decoder frameworks. The model proposed uses a neural network to store knowledge where it acknowledged that it is not possible to store all data present in the real world. The model uses sequence to sequence learning and end to end format is used for answers so as to get fluid responses . Context from conclusion decoder is used to create supplementary decoder states using attention mechanism and closeness of encoders output sequence is compared to separate outputs of conclusion and supplement decoders output sequence . Resulting answers match question . Experimentation done on datasets including “Love Advice” and “Arts & Humanities” categories observes more accurate results than other tested baseline models.

In [29], Proposed attention based model generates caption for photos given as input and uses pc vision strategies to grasp content of image and linguistic communication process for understanding image content into words. Experimentation is done on MS COCO , Flickr30k and Flickr8k dataset . Model uses a backpropagation technique to train and automatically fixes salient objects in generated output. Soft and hard attention mechanisms are used for generating image captions and BLUE and METOR metrics are used to evaluate state of the benchmark datasets.

In[30], a multi-turn open-domain chatbot trained end-to-end on data mined and filtered from public domain social media conversations is proposed . This neural network used has 2.6 billion parameters and minimizes perplexity of the next token. A human evaluation metric called Sensibleness and Specificity Average (SSA),is also proposed. It uses seq2seq model with Evolved Transformer (ET) and is trained on multi-turn conversations with the input sequence .

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

III. BACKGROUND

3.1 Natural Language Processing

It is a subfield of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language, in particular how to program computers to process and analyze large amounts of natural language data.

3.2 GPT-2

Generative Pre-trained Transformer 2 is an open-source AI. created by OpenAI in February2019. GPT-2 translates text, answers questions, summarizes passages, and generates text output on a level that, while sometimes indistinguishable from that of humans,[6] can become repetitive or nonsensical when generating long passages. The architecture implements a deep neural network, specifically a transformer model,[9] which uses attention in place of previous recurrence- and convolution-based architectures.[10][11] Attention mechanisms allow the model to selectively focus on segments of input text it predicts to be the most relevant.[12][13] This model allows for greatly increased parallelization, and outperforms previous benchmarks for RNN/CNN/LSTM-based models

IV. PROPOSED SOLUTION

The proposed system has been trained on a custom dataset which includes abstracts of over 6000 research papers. The main aim was to build a content writer which can produce high-quality content while also being creative. The proposed system can be useful to develop content for readers on a daily basis. It generates articles based on a snippet of text, a sentence or even a passage from an article given as input by the user. It allows the user to define the length of text and degree of innovation in the text to be produced. We employ a transformer-decoder language model also known as GPT-2 which has 124 million parameters. The model is fine-tuned till desired results were achieved in the produced content. The recommended temperature lies between 0.7 to 0.9 which produced good results.

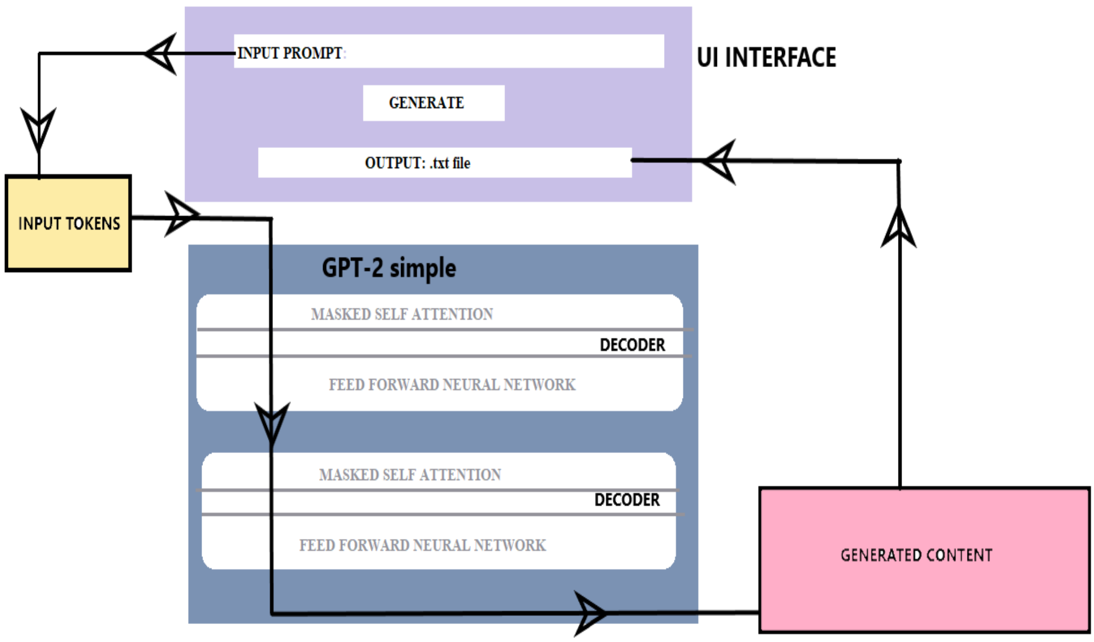


Figure 1: Proposed Architecture

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Execution steps are as follows:

Step 1: Open the web interface to give the input i.e. title, text snippet.

Step 2: Click on the generate button.

Step3: Now the model will take the input and convert them into tokens and passes through transformer decoder block and produces new content

Step 4: Generates content which is available in .txt format that can be downloaded.

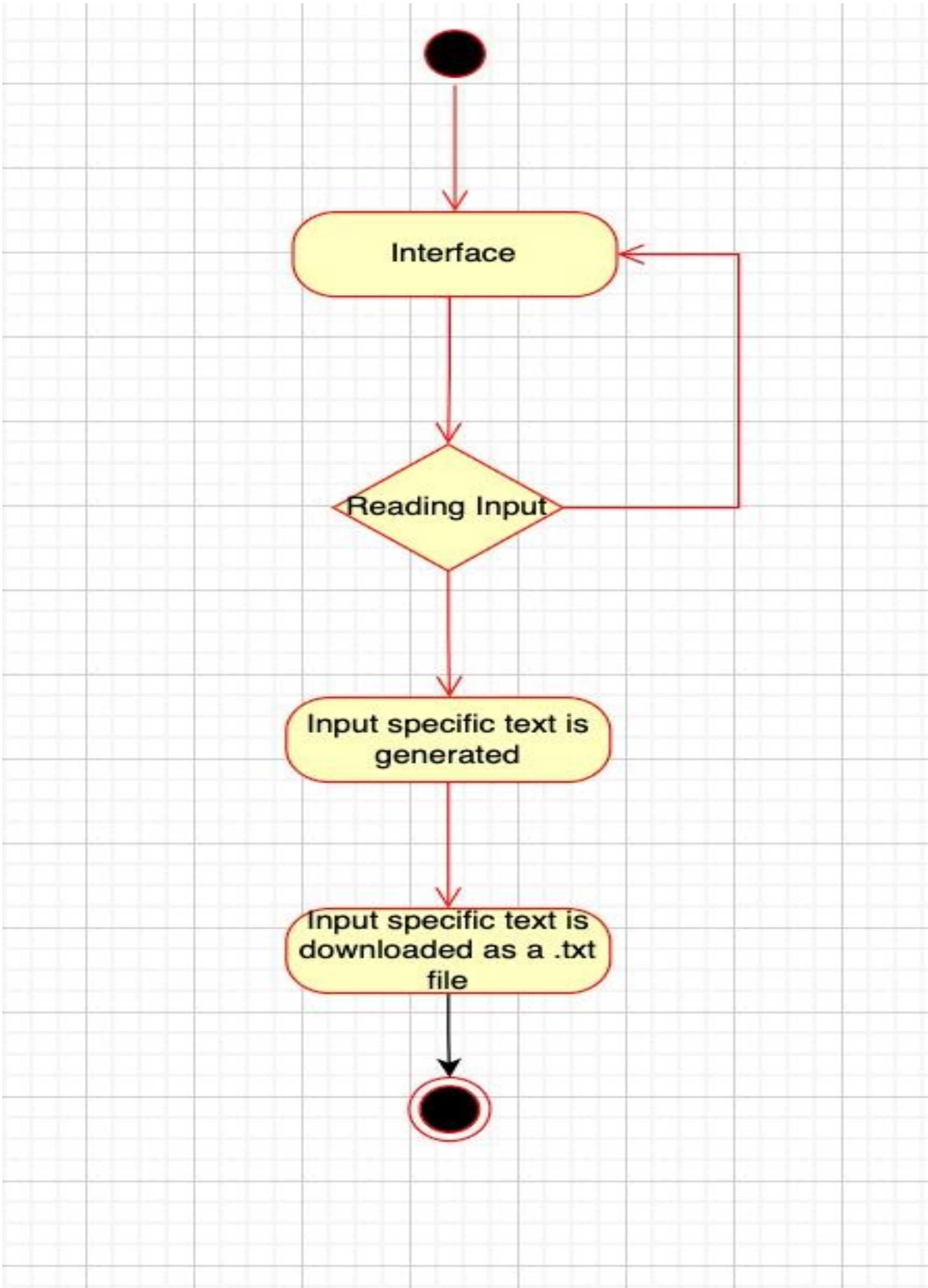


Figure 2: Execution Flow

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

V. RESULTS

Figure 3: UI High Performance Content Writer.

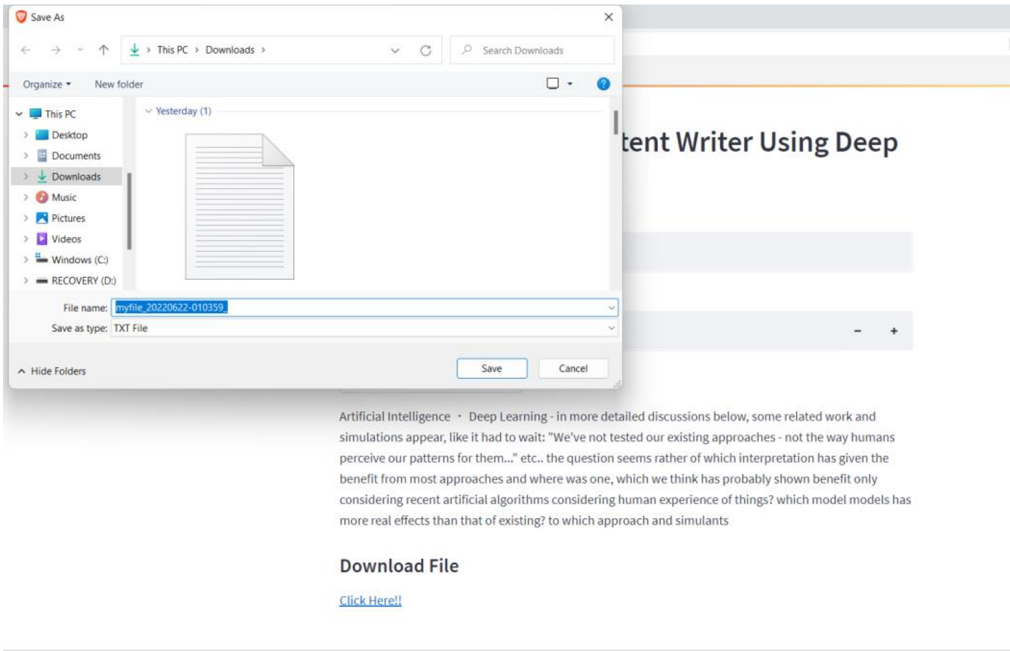
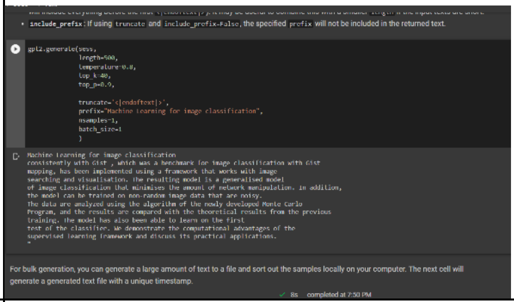
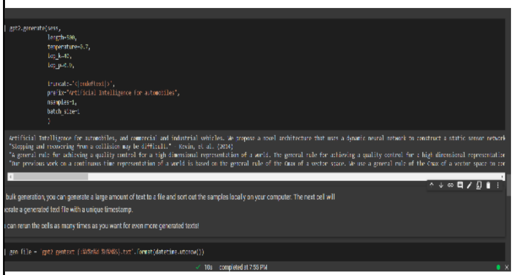


Figure 4: Generated content can be downloaded by the user through the interface and is available in a .txt file.

Table 1: Results

Test case ID	Test case name	Purpose	Test Case	Output
1	Content generation test1	To check if the model gives the content related to the title	User enter the title of the article to be generated	Model generates content related to the title 
2.	Content generation test2	To check whether the model generates the specified length in accordance	User specifies the length of the article to be generated	Model generates article according to specified length 
3.	Content generation test3	To check if the model produces content that is not repetitive	User can read the content generated by the model	The content generates by the model is not ambiguous
4	Content generation test4	To check if the provided information is factual and not gibberish	User can read the content generated by the model	The content produced by the model is trained on a 70mb factual data

VI. CONCLUSION

Applications of artificial intelligence to generate content is the next big thing that will create its own place in content creation. Content creation still represents a standing challenge for deep-learning NLP. Even more so this task is applied to a domain-specific corpus that are different from the pre-training. highly technical, or contains low amount of training materials. Nevertheless, we have here illustrated that the text-to-text, multi-loss training strategy could be used to fine-tune a pre-trained language model. such as GPT-2 for content generation. The result is interpretable and reasonable, even though it is not near human-level performance. We think that our model could benefit from further training .This should make the model more accurate its ability. As AI grows more sophisticated, figuring out how to enable the good uses without the bad ones will be one of our biggest challenge. At the end of the day, AI should be used as a tool to improve and accompany the content writing process, not be the sole source of copy.

ACKNOWLEDGMENT

We thank CMR Technical Campus for supporting this paper titled "Malware analysis with Machine learning: Classifying malware based on PE Header " ,which provided good facilities and support to accomplish our work. Sincerely thank our Chairman, Director, Deans, Head Of the Department, Department Of Computer Science and Engineering, Guide and Teaching and Non- Teaching faculty members for giving valuable suggestions and guidance in every aspect of our work.

REFERENCES

[1] Y. Qu, P. Liu, W. Song, L. Liu and M. Cheng, "A Text Generation and Prediction System: Pre-training on New Corpora Using BERT and GPT-2," *2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, 2020, pp. 323-326, doi: 10.1109/ICEIEC49280.2020.9152352.

[2] Kieuvongngam, Virapat, Bowen Tan and Yiming Niu. “Automatic Text Summarization of COVID-19 Medical Research Articles using BERT and GPT-2.” *ArXiv* abs/2006.01997 (2020): n. page.

[3] Bartoli, Alberto and Eric Medvet. “Exploring the Potential of GPT-2 for.” (2020).

[4] Radford, Alec, Jeff Wu, Rewon Child, David Luan, Dario Amodei and Ilya Sutskever. “Language Models are Unsupervised Multitask Learners.” (2019).

[5] Dathathri, Sumanth and Madotto, Andrea and Lan, Janice and Hung, Jane and Frank, Eric and Molino, Piero and Yosinski, Jason and Liu, Rosanne(2019) “Plug and Play Language Models: A Simple Approach to Controlled Text Generation, accessed 16,may,2022, https://arxiv.org/abs/1912.02164

[6] Mehta, Sanket Vaibhav and Rao, Jinfeng and Tay, Yi and Kale, Mihir and Parikh, Ankur P. and Strubell, Emma(2021) “Improving Compositional Generalization with Self-Training for Data-to-Text Generation”, accessed 25,June,2022, https://arxiv.org/abs/2110.08467

[7]. Tianyu Liu, Kexiang Wang, Lei Sha, Baobao Chang, Zhifang Sui “Table- to- Text Generation by Structure- Aware Seq2seq Learning” The Thirty-second AAAI Conference on Artificial Intelligence (AAAI-18) 2018

[8] Schick, Timo and Schütze, Hinrich(2020) “Few-Shot Text Generation with Pattern-Exploiting Training”, https://arxiv.org/abs/2012.11926

[9] Hu, Zhiting and Shi, Haoran and Tan, Bowen and Wang, Wentao and Yang, Zichao and Zhao, Tiancheng and He, Junxian and Qin, Lianhui and Wang, Di and Ma, Xuezhe and Liu, Zhengzhong and Liang, Xiaodan and Zhu, Wangrong and Sachan, Devendra Singh and Xing, Eric P.(2018) “Texar: A Modularized, Versatile, and Extensible Toolkit for Text Generation”, arXiv:1809.00794v2

[10] Shleifer, Sam and Rush, Alexander M.(2020) “Pre-trained Summarization Distillation”, https://arxiv.org/abs/2010.13002v2

[11] N. Fatima, A. S. Imran, Z. Kastrati, S. M. Daudpota and A. Soomro, "A Systematic Literature Review on Text Generation Using Deep Neural Network Models," in *IEEE Access*, vol. 10, pp. 53490-53503, 2022, doi: 10.1109/ACCESS.2022.3174108.

[12] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," in IEEE Transactions on Signal Processing, vol. 45, no. 11, pp. 2673-2681, Nov. 1997, doi: 10.1109/78.650093.

[13] Touseef Iqbal, Shaima Qureshi(2022)"The survey: Text generation models in deep learning", Journal of King Saud University - Computer and Information Sciences,Volume 34, Issue 6, Part A, Pages 2515-2528,ISSN 1319-1578,https://doi.org/10.1016/j.jksuci.2020.04.001.(https://www.sciencedirect.com/science/article/pii/S1319157820303360)

[14] Dai, Zihang and Yang, Zhilin and Yang, Yiming and Carbonell, Jaime and Le, Quoc V. and Salakhutdinov, Ruslan(2019) "Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context", https://arxiv.org/abs/1901.02860

[15] Zhang, Wen and Feng, Yang and Meng, Fandong and You, Di and Liu, Qun(2019) "Bridging the Gap between Training and Inference for Neural Machine Translation", https://arxiv.org/abs/1906.02448v2

[16] Devlin, Jacob and Chang, Ming-Wei and Lee, Kenton and Toutanova, Kristina(2018) "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", https://arxiv.org/abs/1810.04805v2

[17] Vinyals, Oriol and Le, Quoc(2015) "A Neural Conversational Model", https://arxiv.org/abs/1506.05869v3

[18] Radford, Alec and Karthik Narasimhan. "Improving Language Understanding by Generative Pre-Training." (2018).

[19] Gao, Jianfeng and Galley, Michel and Li, Lihong(2018) "Neural Approaches to Conversational AI", https://arxiv.org/abs/1809.08267v3

[20] Zhilin Yang and Zihang Dai and Yiming Yang and Jaime Carbonell and Ruslan Salakhutdinov and Quoc V. Le(2019), "XLNet: Generalized Autoregressive Pretraining for Language Understanding", https://doi.org/10.48550/arXiv.1906.08237

[21] Lan, Zhenzhong and Chen, Mingda and Goodman, Sebastian and Gimpel, Kevin and Sharma, Piyush and Soricut, Radu(2019) "ALBERT: A Lite BERT for Self-supervised Learning of Language Representations", https://arxiv.org/abs/1909.11942v6

[22]. Xia, Rui and Ding, Zixiang(2019) "Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts", https://arxiv.org/abs/1906.01267v1

[23]. Anja Belz,"Automatic Generation of Weather Forecast Texts Using Comprehensive Probabilistic Generation-Space Models", Natural Language Engineering 1 (1): 1-26. Printed in the United Kindom

[24]. Lebert,R; Grangier,D; and Auli,M. 2016.,"Neural text generation from structured data with application to the biography domain", arXiv preprint arXiv:1603.07771

[25]. Longxu Dou , Guanghui Qin , Jinpeng Wang , Jin-Ge Yao3 , and Chin-Yew Lin3," Data2Text Studio: Automated Text Generation from Structured Data", Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (System Demonstrations), pages 13–18

[26]. Ratish Puduppully and Li Dong and Mirella Lapata "Data-to-text Generation with Entity Modeling", Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages.

[27]. Adwait Ratnaparkhi,"Trainable approaches to surface natural language generation and their application to conversational dialog systems", 2002 Computer Speech & Language, 16(3): 435-455.

[28]. J. Yin, X. Jiang, Z. Lu, L. Shang, H. Li, and X. Li, "Neural generative question answering," in Proc. IJCAI,2016, pp. 2972–2978.

[29]. Kelvin Xu, Jimmy Lei Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhutdinov, Richard S. Zemel, Yoshua Bengio, "Show, Attend and Tell: Neural Image Caption Generation with Visual Attention", arXiv: 1502.03044v3 [cs.LG] 19 Apr 2022.

[30] Adiwardana, Daniel and Luong, Minh-Thang and So, David R. and Hall, Jamie and Fiedel, Noah and Thoppilan, Romal and Yang, Zi and Kulshreshtha, Apoorv and Nemade, Gaurav and Lu, Yifeng and Le, Quoc V.(2020) "Towards a Human-like Open-Domain Chatbot", https://doi.org/10.48550/arXiv.2001.09977

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Paridhi Wahii , Abisha Winslet Beri , Suchith Adepu conducted research on text-generation algorithms and procedures . They explored and built a custom dataset on which a fine-tuned simple gpt-2 model is trained.

DR. T.Subha Mastan Rao guided the project as an academic project mentor.