BengaliGPT: An Instruction Following LLaMA Model for Bengali

Preprint · June 2023

DOI: 10.13140/RG.2.2.16116.26247

7 authors, including:

Arghyadeep Sen
KIIT University
8 PUBLICATIONS **27** CITATIONS
SEE PROFILE

Shashikanta Sahoo
Biju Patnaik University of Technology
5 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Shantipriya Parida
Silo AI
106 PUBLICATIONS **597** CITATIONS
SEE PROFILE

Satya Ranjan Dash
KIIT University
170 PUBLICATIONS **808** CITATIONS
SEE PROFILE

BengaliGPT: An Instruction Following LLaMA Model for Bengali

Guneet Singh Kohli
Thapar University, India
guneetsk99@gmail.com

Arghyadeep Sen
KIIT University, India
senarghyadeep@gmail.com

Sambit Sekhar
Odia Generative AI, Bhubaneswar, India
sambitskhr@gmail.com

Shashikanta Sahoo
Govt College Kalahandi, India
shashisahoo123@gmail.com

Shantipriya Parida
Silo AI, Helsinki, Finland
shantipriya.parida@silo.ai

Satya Ranjan Dash
KIIT University, Bhubaneswar, India
sdashfca@kiit.ac.in

Ondřej Bojar
Charles University, MFF ÚFAL, Czech Republic
bojar@ufal.mff.cuni.cz

Abstract

Generative artificial intelligence and Large language models (LLMs) are significantly impacting the AI community. However, these technologies are still limited to a few languages which is a great concern. In a multilingual country like India, where the majority of the population communicates in their native languages other than English, the need for LLM models adapted to regional languages becomes crucial.

In this paper, we propose an instruction following the BengaliGPT model for the Bengali language; the seventh most spoken language in the world. Our dataset includes 252K Bengali instruction sets and these are translated from several open-source resources such as Alpaca-COT, Dolly, GPT teacher instruct, GPT teacher roleplay, and others. We have adopted LoRA for model building and LLaMA for fine-tuning. As a result, BengaliGPT shows major improvements in answering simple instruction-based queries and the model is able to provide almost accurate Bengali output for such queries. Our BengaliGPT will be available freely for research and non-commercial purposes.

1 Introduction

In early childhood, a person develops the ability to understand, express, and communicate in a specific language and this ability evolves within the entire lifetime. In the context of machines, there is no such ability unless an Artificial Intelligence-driven algorithm is applied to imitate such natural ability of a person (Ouyang et al., 2022). Hence, this context became a challenge for researchers so that machines can read, write, and communicate with a person like a natural human being. Language Modelling is considered as an approach to develop a model that can predict probabilities of tokens based on the likelihood of word sequences (Lampinen et al., 2022). In current research on Natural Language Processing (NLP), LM became a major focus for experimentation, model development, fine-tuning, language-based architectural, and algorithm studies.

Large Language Models (LLMs) have accelerated progress of NLP research to its peak and there are several products that are launched for millions of users such as Google Search Engine (after BERT being employed)[1], ChatGPT[2], and coding assistant Copilot (Chen et al., 2022). Language understanding could be achievable when memorization (Mialon et al., 2023) is combined with prompting at conditional or unconditional grounds improving Human-Computer Interactions.

Some major limitations of LLMs are identified that are preventing widespread deployment. Large Language Models can provide predictions that are seemingly probable or unrelated to the context that is often denoted as hallucinations (Welleck et al., 2019). Therefore, mistakes could be avoided in the context of arithmetic or reasoning. However, many LLMs can potentially resolve problems with few-shot learning capabilities when a certain scale is achieved (Mialon et al., 2023). The size and volume of data that are necessary to train LLMs could seem impractical and according to (Scialom et al., 2022), continual learning is still an open research question.

In this paper, we propose a BengaliGPT model by fine-tuning the Bengali instruction set following Low-Rank Adaptation (LoRA). Also, we propose a benchmarked dataset for evaluation which can be useful for Bengali LLM.

2 Literature Survey

Researchers have cited that Pre-trained Language Models and scaling of their model size and data size can lead to improved model capabilities for downstream tasks (Kaplan et al., 2020). For example, large-sized Pre-trained Language Models (PLMs) such as 330M-parameter BERT and 1.5B-parameter GPT-2 show hidden abilities or emergent abilities to solve some complex tasks (Wei et al., 2022). GPT-3 could solve few-shot learn-

[1] https://blog.google/products/search/search-language-understanding-bert/

[2] https://openai.com/blog/chatgpt

ing tasks using in-context learning. Hence, the researchers have coined the term "Large Language Models" to refer to these large-sized PLMs (Shanahan, 2022). One of the significant applications of LLM is ChatGPT, where LLMs can have conversations with real-world people.

LLMs are considered a revolutionary step towards combining human interaction with Artificial Intelligence (AI) interfaces with prompting. The research areas of AI are becoming more focused on progressing with the development of LLMs. LLM can serve as the solver of language tasks and in NLP fields, LLM would be used as a primary tool (Huang et al., 2023). In the context of Information Retrieval (IR), ChatGPT is challenging traditional search engines; however, New Bing[3] provided an AI-driven search engine that works with LLMs. In the context of Computer Vision (CV), new language models should be able to serve with multimodal inputs and GPT-4 supports multimodal inputs with the integration of visual information (Driess et al., 2023). Therefore, developing LLM with multimodal information processing helps build the NLP research base empowered with Copilot as a coding assistant in Microsoft 365 and plugins in ChatGPT for special functions (Wu et al., 2023).

3 Focused Language

The Bengali language also named Bangla is an Indo-Aryan language that is native to the Bengal region of South Asia. It is the official language of Bangladesh and the Indian state of West Bengal and the seventh most spoken language in the world (Sen et al., 2022).

Our research mainly focuses on using Bengali as the primary language to interact with the LLM. Bengali is considered a low-resource language and the development of LLM in this language based on instruction sets would be useful for developing chatbots and solving few-shot learning tasks.

4 Dataset Details

The dataset contains a 252K Bengali instruction set. The instruction set is translated data (English-to-Bengali) from open-source resources, resulting in good Bengali instruction understanding and response generation capabilities.

The statistics of the dataset are shown in Table 1.

5 Model Building

For model building we adopted, Low-Rank Adaptation (LoRA) which freezes the pre-trained model weights and injects trainable rank decomposition matrices into each layer of the Transformer architecture (Hu et al., 2021). We used Large Language Model Meta AI (LLaMA) (Touvron et al., 2023) as the foundation model for fine-tuning. Due to the smaller size, LLaMA requires fewer computing resources, and we used LLaMA 7B for fine-tuning, which is trained on one trillion tokens with a majority of data in English.

[3]https://www.bing.com/new

Dataset	Size
Alpaca (Taori et al., 2023a)	60,402
Dolly	54,456
GPT teacher	9,111
GPT teacher instruct	9,987
Hard code Q&A	18,194

Table 1: Details of the data used in the instruction fine-tuning stage.

We followed a methodology similar to that employed in Stanford Alpaca (Taori et al., 2023b) to implement self-guided fine-tuning for training the instruction-following model. Each instance comprises an instruction and a corresponding output.

{### *Instruction:*}
{*instruction*}
{### *Response: output*}

{### *Instruction:*}
{*instruction*}
{*Input:*}
{### *Response: output*}

The instruction is fed into the model, and the model is prompted to generate the output in an autoregressive manner (Zhao et al., 2023). This process is akin to the task of casual language modeling (Feder et al., 2021). For self-instructed fine-tuning, we utilize a prompt template derived from Stanford Alpaca, also applied during the inference phase.

6 Experimental Setting

The experimental setting for this research paper involved training a language model using an Nvidia A100 PCIE GPU with 40 GB of memory. The model was trained for a total of five epochs, which took approximately four days to complete. Several hyperparameters were utilized during the training process to optimize the model's performance. A batch size of 128 was employed, along with a learning rate of 3e−4. To prevent overfitting, a weight decay of 0.001 was applied. The training process also incorporated a warmup rate of 0.1 to gradually increase the learning rate. The learning rate scheduler followed a linear function. The model architecture utilized a Lora r of 16 and targeted specific modules, including q_proj, k_proj, v_proj,

and o_proj. Additionally, a cutoff length of 256 was used to limit the input length during training. These experimental settings were carefully selected to ensure an optimal balance between computational resources and model performance.

7 Training

We trained the model on a single GPT for 5 epochs. The training hyperparameters are shown in Table 2.

Hyper Parameter	Value
Batch Size	128
Learning Rate	$3e^{-4}$
Epochs	5
Cutoff Length	256
Weight_Decay	0.001
Warmup_Rate	0.1
LR_Scheduler	linear
Lora r	16
Lora Target Modules	(q_proj, k_proj, v_proj, o_proj)

Table 2: Training Hyperparameters.

The training and evaluation loss are shown in Figure 1 and Figure 6.

Figure 1: Training loss.

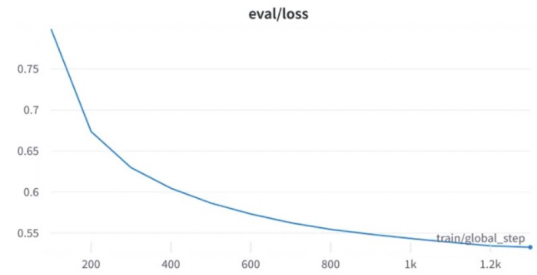

Figure 2: Evaluation loss.

8 Inference

8.1 Text Generation Setup

The decoding process of LLMs plays a critical role in determining the quality and diversity of the generated text. In our experiments, we use the following decoding hyperparameters:

- *Size of the context*: We establish the context size as 2048, determining the maximum number of tokens that the model can take into account simultaneously during the text generation process.
- *Maximum sequence length*: We impose a constraint on the generated sequence length, limiting it to 512 tokens to ensure that the outputs remain focused and closely related to the input prompt.
- *Temperature*: We set the temperature to 0.2, regulating the level of randomness in the sampling process. Lower values make the model produce more focused and deterministic outputs, while higher values introduce greater diversity at the expense of coherence.
- *Top-k sampling*: For each step, we adopt Top-k sampling with a value of k = 40, whereby the model selects the subsequent token from the top 40 most probable options. This introduces an element of randomness and diversity in the generated text.

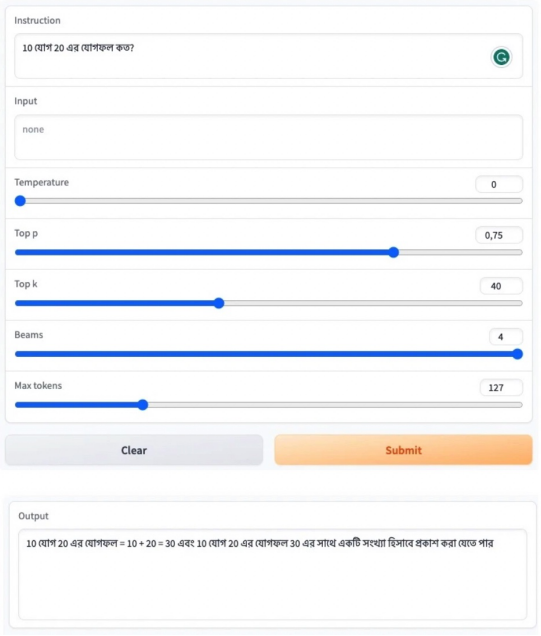

Figure 3: Sample Inference 1. The question is, "What is the sum of 10 plus 20?" The answer is "The sum of 10 plus 20 = 10 + 20 = 30 and the sum of 10 plus 20 can be expressed as a number with 30"

8.2 Automatic Evaluation

We evaluated the BengaliGPT model on natural language generation (NLG) tasks including text summarization. We used Rouge (LIN, 2004), BLEU (Post, 2018), and BertScore (Zhang et al., 2019) automatic evaluation as shown in Table 3.

Figure 4: Sample Inference 2. The question is, "What are the benefits of eating an apple a day?" The answer is "Benefits of eating an apple a day Apples are a healthy and wholesome food to eat."

Figure 5: Sample Inference 3. The question is, "What is the primary source of energy that causes evaporation of water from the surface of a body of water?" input is, "text": ["solar radiation", "conduction by plants", "heat from surrounding land mass", "convection currents in water"], "label": ["A", "B", "C", "D"] The answer is "Solar radiation by plants is a primary source of energy that causes evaporation of water from the surface of water bodies."

This section presents a comprehensive analysis of the results obtained from evaluating our newly developed Bengali LLM. The evaluation focused on two specific tasks: Natural Language Generation (NLG) and Summarizing and Rephrasing. We aimed to assess the model's performance in generating coherent and contextually appropriate text in the Bengali language. Furthermore, we compared the results of our model with those of ChatGPT, a widely used multilingual LLM, to gauge the performance and adaptability of our model. For the NLG task, we designed a sample dataset consisting of 100 elements. The dataset encompassed a range of tasks, including sentence generation from three given words and paragraph generation on specific themes. We employed two famous metrics as a performance measure: BERT Score and Rouge-Bleu Score. The BERT Score measures the similarity between the generated text and the ground truth, considering precision and recall. Our Bengali LLM achieved a precision score of 0.88 and a recall score of 0.91, indicating its ability to generate text that closely aligns with the desired output. This demonstrates that our model generates accurate and contextually appropriate sentences and paragraphs. The Rouge and Bleu scores provide a comprehensive evaluation of the generated text's quality by assessing various aspects such as sentence-level overlap (rouge1), bi-gram overlap (rouge2), and longest common subsequence (rougeL). Our Bengali LLM achieved an F1 score of 0.89, indicating a balance between precision and recall. Additionally, the RougeLSum score of 0.34 demonstrates the model's ability to generate text consistent with the reference summaries. Finally, the Bleu Score of 34.3 further emphasizes the model's competence in generating text similar to the ground truth. In addition to the NLG task, we also evaluated the performance of our Bengali LLM in the Summarize and Rephrase task. Here, we focused on generating headlines from sentences and rephrasing news articles. We had access to human-annotated answers for this evaluation, which served as ground truths

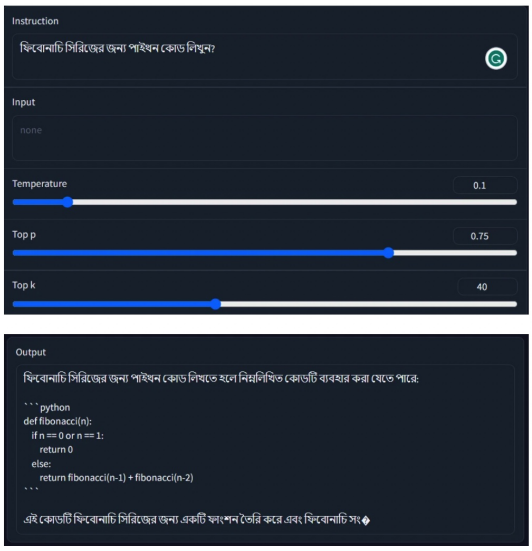

Figure 6: Sample Inference 4. The question is, "Write Python code for the Fibonacci Series". The answer "The following code can be used to write python code for the Fibonacci Series [python code]"

for our model. The results obtained for the Summarize and Rephrase task were promising, with our Bengali LLM achieving a precision score of 0.88 and a recall score of 0.89. This indicates that the model effectively captures the critical information from the source sentences and generates concise and accurate headlines and rephrased news articles. The Rouge and Bleu scores for the Summarize and Rephrase task further reinforce the model's proficiency. The F1 score of 0.89 reflects the balance between precision and recall, while the RougeLSum score of 0.37 demonstrates the model's ability to generate summaries that align well with the reference summaries. Moreover, the Bleu Score of 35.5 highlights the model's capability to generate text that closely resembles the ground truth. Overall, the results obtained from the evaluation of our Bengali LLM indicate its strong performance in the NLG and Summarize and Rephrase tasks. The model showcases remarkable adaptability to multilingual data, as demonstrated by its ability to generate text in the Bengali language, which is the focus of this evaluation. The positive response from our Bengali LLM underscores its potential as a foundational multilingual LLM. The evaluation strategy adopted for this study holds significant implications for future evaluations of low-resource Large Language Models. We can ensure reliable assessments of these models' capabilities across various tasks and languages by developing and applying rigorous evaluation methodologies. This research sets a precedent for evaluating and benchmarking multilingual LLMs, thereby contributing to advancing the field and developing more effective and versatile language models in the future.

8.3 Human Evaluation

We have used different types of instruction sets for testing out the model performance. Based on the model outcomes, we have employed a native Bengali speaker to assess the model outcomes. The instructions were Natural Language Generation, Sentence generation using three given words, basic code writing, text summary writing and content rephrasing (refer to Table 5). We have also selected different questions to check the model's ability to generate answers, ranging from mathematical questions, reasoning questions, code generation, etc. The statistics of the dataset are shown in Table 4. We have performed human evaluation over mathematical questions, philosophical and logical fallacies context. The model performance was not up to the mark; there were repetitive sentence segments, meaningless outputs and mistakes in some mathematical outputs.

As per native speaker's opinion, the score for appropriateness was 2/10, content was 4/10, readability was 5/10, grammar of the output was 2/10 and relevance was 3/10. These metrics were considered from the paper cited as (Lin and Chen, 2023).

Only correct outcome we have obtained from mathematical instructions can be shown as:

−40 ÷ (−8) এর ভাগফল নির্ণয় কর। −40 ÷ (−8) = 5

We have instructed the model to divide -40 by -8 and the output is given as 5. The outcome is correct in here.

5.04/0.7 এর ভাগফল নির্ণয় কর 5.04/0.7 = 7.04

We have instructed the model to divide 5.04 by 0.7 and the output is almost correct here. The answer should be 7.2.

ম্যাডি 7 দিনের মধ্যে মোট 56 মাইল তার সাইকেল চালাবে। তিনি প্রতিদিন একই সংখ্যক মাইল চালাবেন। ম্যাডি প্রতিদিন মোট কত মাইল যাত্রা করবে? ম্যাডি প্রতিদিন মোট 56/7 = 8 মাইল চালাবে।

We have given the model a situation where Maddy would ride her bicycle for 7 days and she will cover total of 56 mile during her journey. We have asked the model to calculate how far she will go everyday considering she rode same distance each day. The output comes out to be 8 mile which is precisely correct.

প্রদত্ত সংখ্যাগুলি থেকে $x + 2.7 = 8.4$ সমীকরণের একটি সমাধান চিহ্নিত করুন। $x + 2.7 = 8.4$, এই সমীকরণের সমাধান হলঃ $x = 8.4 - 2.7$; $x = 5.7$; তাই, সংখ্যা x এর সমাধান হলঃ 5.7

We have asked the model to identify a value from the equation $x + 2.7 = 8.4$ and the model correctly identified the outcome to be 5.7.

We can reach to the conclusion that current model can answer only Natural Language Generation, Sentence generation using three given words, basic code writing, text summary writing and content rephrasing - at a very basic level. The model is unable to perform well with critical reasoning and mathematical computations. We plan to release this benchmark dataset for LLM evaluation, particularly for the Bengali language considering the unavailability of such benchmark dataset for Indic languages.

9 Conclusion and Future Plan

The LLMs have the potential to revolutionize the field of AI, enabling machines to better understand and interact with humans in a natural language setting.

The plan includes:

- Adapting QLoRA (Dettmers et al., 2023) for fine-tuning.
- Fine-tuning with a more domain-specific Bengali instruction set.
- Pre-train Bengali LLM model,
- Chatbot development supporting Bengali.

Tasks	Bert Score		Rouge					
	Precision	**Recall**	**F1**	**rouge1**	**rouge2**	**rougeL**	**rougeLSum**	**Bleu Score**
NLG	0.88	0.91	0.89	0.43	0.15	0.35	0.34	34.3
Summarize and Rephrase	0.88	0.89	0.89	0.4	0.23	0.36	0.37	35.5

Table 3: NLG Metrics Comparison

Dataset	**Size** (number of questions)
TruthfulQA	20
LogiQa	20
MMLU Logical Fallacies	80
MMLU Philosophy	30
MMLU Jurisprudence	30

Table 4: Manual Evaluation Dataset Statistics

Limitations

Although the current model is able to accept Bengali input text and generate answers in Bengali, it still fails to answer questions related to general knowledge about India due to a lack of domain knowledge. Also, it fails in critical reasoning as per evaluation.

Ethics Statement

We do not envisage any ethical concerns. The dataset does not contain any personal, or personally identifiable, information, the source data is already open source, and there are no risks or harm associated with its usage.

References

Wenhu Chen, Xueguang Ma, Xinyi Wang, and William W Cohen. 2022. Program of thoughts prompting: Disentangling computation from reasoning for numerical reasoning tasks. *arXiv preprint arXiv:2211.12588*.

Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. 2023. Qlora: Efficient finetuning of quantized llms. *arXiv preprint arXiv:2305.14314*.

Danny Driess, Fei Xia, Mehdi SM Sajjadi, Corey Lynch, Aakanksha Chowdhery, Brian Ichter, Ayzaan Wahid, Jonathan Tompson, Quan Vuong, Tianhe Yu, et al. 2023. Palm-e: An embodied multimodal language model. *arXiv preprint arXiv:2303.03378*.

Amir Feder, Nadav Oved, Uri Shalit, and Roi Reichart. 2021. Causalm: Causal model explanation through counterfactual language models. *Computational Linguistics*, 47(2):333–386.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.

Shaohan Huang, Li Dong, Wenhui Wang, Yaru Hao, Saksham Singhal, Shuming Ma, Tengchao Lv, Lei Cui, Owais Khan Mohammed, Qiang Liu, et al. 2023. Language is not all you need: Aligning perception with language models. *arXiv preprint arXiv:2302.14045*.

Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. 2020. Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*.

Andrew K Lampinen, Ishita Dasgupta, Stephanie CY Chan, Kory Matthewson, Michael Henry Tessler, Antonia Creswell, James L McClelland, Jane X Wang, and Felix Hill. 2022. Can language models learn from explanations in context? *arXiv preprint arXiv:2204.02329*.

C-Y LIN. 2004. Rouge: A package for automatic evaluation of summaries. In *Proc. of Workshop on Text Summarization Branches Out, Post Conference Workshop of ACL 2004*.

Yen-Ting Lin and Yun-Nung Chen. 2023. Llm-eval: Unified multi-dimensional automatic evaluation for open-domain conversations with large language models. *arXiv preprint arXiv:2305.13711*.

Grégoire Mialon, Roberto Dessì, Maria Lomeli, Christoforos Nalmpantis, Ram Pasunuru, Roberta Raileanu, Baptiste Rozière, Timo Schick, Jane Dwivedi-Yu, Asli Celikyilmaz, et al. 2023. Augmented language models: a survey. *arXiv preprint arXiv:2302.07842*.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744.

Matt Post. 2018. A call for clarity in reporting BLEU scores. In *Proceedings of the Third Conference on Machine Translation: Research Papers*, pages 186–191. Association for Computational Linguistics.

Thomas Scialom, Tuhin Chakrabarty, and Smaranda Muresan. 2022. Continual-t0: Progressively instructing 50+ tasks to language models without forgetting. *arXiv preprint arXiv:2205.12393*.

Ovishake Sen, Mohtasim Fuad, Md Nazrul Islam, Jakaria Rabbi, Mehedi Masud, Md Kamrul Hasan, Md Abdul Awal, Awal Ahmed Fime, Md Tahmid Hasan Fuad, Delowar Sikder, et al. 2022. Bangla natural language processing: A comprehensive analysis of classical, machine learning, and deep learning based methods. *IEEE Access*.

Murray Shanahan. 2022. Talking about large language models. *arXiv preprint arXiv:2212.03551*.

Rohan Taori, Ishaan Gulrajani, Tianyi Zhang, Yann Dubois, Xuechen Li, Carlos Guestrin, Percy Liang, and Tatsunori B Hashimoto. 2023a. Alpaca: A strong, replicable instruction-following model. *Stanford Center for Research on Foundation Models. https://crfm. stanford. edu/2023/03/13/alpaca. html*, 3(6):7.

Rohan Taori, Ishaan Gulrajani, Tianyi Zhang, Yann Dubois, Xuechen Li, Carlos Guestrin, Percy Liang, and Tatsunori B Hashimoto. 2023b. Stanford alpaca: An instruction-following llama model.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, et al. 2022. Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*.

Sean Welleck, Ilia Kulikov, Stephen Roller, Emily Dinan, Kyunghyun Cho, and Jason Weston. 2019. Neural text generation with unlikelihood training. *arXiv preprint arXiv:1908.04319*.

Chenfei Wu, Shengming Yin, Weizhen Qi, Xiaodong Wang, Zecheng Tang, and Nan Duan. 2023. Visual chatgpt: Talking, drawing and editing with visual foundation models. *arXiv preprint arXiv:2303.04671*.

Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. 2019. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*.

Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, Yifan Du, Chen Yang, Yushuo Chen, Zhipeng Chen, Jinhao Jiang, Ruiyang Ren, Yifan Li, Xinyu Tang, Zikang Liu, Peiyu Liu, Jian-Yun Nie, and Ji-Rong Wen. 2023. A survey of large language models. *arXiv preprint arXiv:2303.18223*.

A Appendix

Tasks	Instruction	BengaliGPT Output
Natural Language Generation	আমি সম্প্রতি অনলাইনে একজোড়া হেডফোন কিনেছি এবং আমি তাদের গুণমান এবং কার্যকারিতা দেখে খুব মুগ্ধ হয়েছি। আমি এমন একটি পণ্যের পর্যালোচনা লিখতে চাই যা আমার ইতিবাচক অভিজ্ঞতাকে প্রতিফলিত করে এবং অন্যান্য সম্ভাব্য ক্রেতাদের একটি জ্ঞাত সিদ্ধান্ত নিতে সাহায্য করে। আমি কীভাবে গঠন করতে পারি? আমার পর্যালোচনা এবং কোন কোন দিকগুলো আমার হাইলাইট করা উচিত?	আমি একটি অনলাইন হেডফোন কিনেছি যা আমার ইতিবাচক অভিজ্ঞতাকে প্রতিফলিত করেছে। এটি একটি অসাধারণ ডিভাইস যা আমার দৈনন্দিন কাজকর্মে সহায়তা করেছে। এটি আমাকে আমার কাজের জন্য সহজে কাজ করতে
Sentence generation from 3 words	এই তিনটি শব্দ ব্যবহার করে একটি বাক্য তৈরি করুন ["বাড়ি", "দূর", "বাজার"]	বাড়ি দূরে বাজারে যাওয়া হয়েছিল।

Code Writing	একটি ফাংশন লিখুন যা একটি স্ট্রিংকে ইনপুট হিসাবে নেয় এবং স্ট্রিংটিকে বিপরীত করে দেয়। নীচের ভাবে ফাংশনটি পরিচিত করা হলঃ
	`python`
	`Copy code`
	`def reverse_string(input_string):`
	`return input_string[::-1]`
	`input_string = "hello"`
	`reversed_string =`
	`reverse_string(input_string)`
	`print(reversed_string)`
		`# Output: olleh`
		
		এখানে "reverse_string" নামক ফাংশনটি প্রদত্ত স্ট্রিংকে রিভার্স করে দেয়। এটি স্ট্রিংটির সর্বশেষ অক্ষর থেকে প্রথম অক্ষরের দিকে যাতে চলে। ফাংশনটি একটি স্ট্রিংকে রিভার্স করে তা প্রিন্ট করে। উদাহরণ হিসাবে, যদি "input_string" এর মান "hello" হয়, তাহলে আউটপুট হবে "olleh"। ব্যবহারঃ আপনি অনুসন্ধানের মধ্যে "input_string" এর মান পরিবর্তন করতে পারেন এবং ফাংশনটি কল করতে পারেন। আপনি পরিবর্তিত "input_string" এর মান নির্দিষ্ট করতে পারেন এবং তারপরে ফাংশনটি কল করতে পারেন। উদাহরণ:
		`python`
		`Copy code`
		`input_string = "hello"`
		`reversed_string =`
		`reverse_string(input_string)`
		`print(reversed_string)`
		`# Output: olleh`
		
		প্রয়োগ: input_string নির্দিষ্ট করুন: input_string = "hello" reverse_string ফাংশন

Summary writing	রাজধানীর মোহাম্মদপুরের একটি বস্তিতে আগুনে দগ্ধ হয়ে চার বছরের এক শিশুর মৃত্যু হয়েছে।	"একটি ছোট শিশু একটি বস্তিতে আগুনে দগ্ধ হয়েছে"
Rephrase using a given context	শরীয়তপুরের নড়িয়া উপজেলায় ভুল চিকিৎসায় এক শিশু মৃত্যুর অভিযোগে এক ক্লিনিকে ভাঙচুর হয়েছে।	"শরীয়তপুরের নড়িয়া উপজেলায় ভুল চিকিৎসায় এক শিশু মৃত্যুর অভিযোগে এক ক্লিনিকে ভাঙচুর হয়েছে। এই অভিযোগের কারণ একটি চিকিৎসা সংস্থার অভিযোগের উপর ভিত্তি করে একটি অভিযোগ চালানো হয়েছে"

Table 5: Some Outputs from our BengaliGPT Model