BioNART: A Biomedical Non-AutoRegressive Transformer for Natural Language Generation

Masaki Asada[1] and **Makoto Miwa**[1,2]
[1]Artificial Intelligence Research Center (AIRC)
National Institute of Advanced Industrial Science and Technology
[2]Toyota Technological Institute, Japan
masaki.asada@aist.go.jp, makoto-miwa@toyota-ti.ac.jp

Abstract

We propose a novel Biomedical domain-specific Non-AutoRegressive Transformer model for natural language generation: BioNART. Our BioNART is based on an encoder-decoder model, and both encoder and decoder are compatible with widely used BERT architecture, which allows benefiting from publicly available pre-trained biomedical language model checkpoints. We performed additional pre-training and fine-tuned BioNART on biomedical summarization and doctor-patient dialogue tasks in English. Experimental results show that our BioNART achieves about 94% of the ROUGE score to the pre-trained autoregressive model while realizing an 18 times faster inference speed on the iCliniq dataset.

1 Introduction

In biomedical natural language processing, several domain-specific Pre-trained Language Models (PLMs) such as BioBERT (Lee et al., 2019), SciBERT (Beltagy et al., 2019) and PubMedBERT (Gu et al., 2021) have been proposed by training on domain-specific text in PubMed/MEDLINE, and these domain-specific models have shown remarkable performance on the in-domain downstream tasks. Similarly, following the success of generative encoder-decoder-based sequence-to-sequence PLMs such as BART (Lewis et al., 2020), T5 (Raffel et al., 2020), and BERT2BERT (Rothe et al., 2020), pre-trained biomedical generative models such as SciFive (Phan et al., 2021) and BioBART (Yuan et al., 2022) have also been proposed, and these models can tackle the generation tasks in the biomedical domain, such as clinical dialogue, biomedical question answering and biomedical text summarization.

The generative models usually adopt an AutoRegressive (AR) strategy to decode the target sequence token-by-token from left to right. AR decoding can perform high-quality inference, but it has the limitation that it cannot be easily parallelized and requires significant time and computational cost for inference. To alleviate this problem, several Non-AutoRegressive (NAR) methods (Lee et al., 2018; Qi et al., 2021) that can simultaneously predict all target tokens have been proposed in the general domain. However, there has been no discussion and verification of NAR methods in the biomedical domain.

This paper proposes a novel biomedical domain-specific NAR model called BioNART. In constructing the biomedical NAR model, we achieve NAR generation with only minimal modifications from the BERT (Devlin et al., 2019)-based encoder-decoder model BERT2BERT (Rothe et al., 2020). This allows our model to benefit from already performed large-scale pre-training by initializing both the encoder and decoder of our model with parameters checkpoints of PLMs such as BioBERT (Lee et al., 2019) and SciBERT (Maheshwari et al., 2021). In addition, we further verify the effect of additional pre-training to fit the parameters from checkpoints to the encoder-decoder NAR. Finally, we perform our pre-training on PubMed/MEDLINE abstracts with permutation language modeling (Yang et al., 2019) task.

Our contributions are summarized as follows:

- We are the first to propose a non-autoregressive language model that enables fast inference for the biomedical domain.
- The encoder and decoder of our non-autoregressive model are compatible with existing BERT architecture to allow the parameter initialization from publicly available biomedical PLMs checkpoints such as BioBERT.
- We perform an additional pre-training that starts with the checkpoints to verify that PLMs checkpoints can be fitted for NAR decoding.

The 22nd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks, pages 369–376

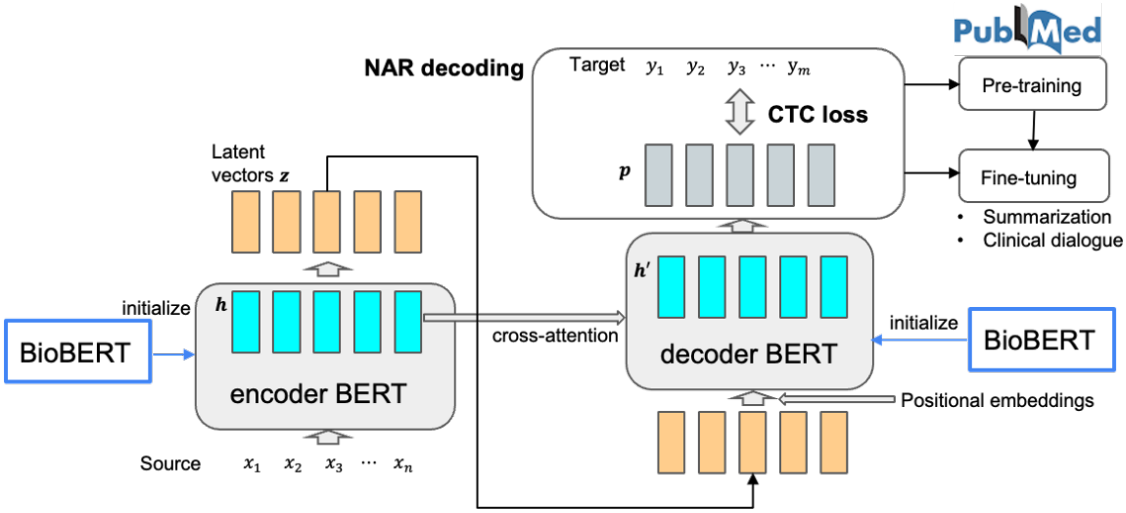

Figure 1: Overview of BioNART

- We evaluate our BioNART models on several biomedical natural language generation tasks and show that leveraging parameter checkpoints and additional pre-training are effective for NAR natural language generation.

2 Method

In this section, we first describe our BioNART model architecture and then explain the objective function used during training and NAR text generation for inference. Finally, we show the parameter initialization of our BioNART. The overview of BioNART is shown in Figure 1.

2.1 BioNART Architecture

Encoder The encoder part of BioNART is exactly the same as the BERT architecture. The source input sentence is split into sub-word tokens w_i by the BERT tokenizer, then converted to word embeddings, positional embeddings, and token type embeddings. These embeddings are fed into the self-attention and feed-forward layers of BERT.

Latent vectors Each token representation of the encoder's final layer $\boldsymbol{h}_i$ is converted into the latent vector $\boldsymbol{z}_i \in \mathbb{R}^{d_z}$ by a single linear layer: $\boldsymbol{z}_i = \boldsymbol{W}_1\boldsymbol{h}_i + \boldsymbol{b}_1$, where $\boldsymbol{W}_1$ and $\boldsymbol{b}_1$ are the weight and bias, and d_z is the dimension of the latent vector. The latent vectors have the same sequence length as the input representations of the decoder.

Decoder In an AR model, the decoder takes the target sentence as input and controls the attention mask so that future tokens are not seen during training. In NAR decoding, the decoder does not take the target sentence as input because target tokens are predicted simultaneously and independently. In the decoder part of BioNART, first, the dimension of the latent vector is matched to the dimension of the hidden representation vectors of the decoder by a single linear layer: $\boldsymbol{h}'_i = \boldsymbol{W}_2\boldsymbol{z}_i + \boldsymbol{b}_2$. Then, these resulting embeddings and positional embeddings are fed into the decoder. The final layer representation from the encoder $\boldsymbol{h}_i$ is also fed into the decoder as a cross-attention. Finally, we calculate the prediction vector $\boldsymbol{p} \in \mathbb{R}^{d_V}$ from the final layer representation of the decoder, where d_V is the vocabulary size.

2.2 Training

Standard AR models are trained with the cross entropy (CE) loss, which compares the prediction $\boldsymbol{p}$ with the target token $\boldsymbol{y}$ at each corresponding position. However, as NAR ignores the dependency in the output token space, it cannot accurately predict token positions. To ease such limitations, several methods (Lee et al., 2018; Ghazvininejad et al., 2019; Gu and Kong, 2021; Li et al., 2022) have been proposed to consider the latent alignments between the target positions. In this work, we adopt Connectionist Temporal Classification (CTC) (Graves et al., 2006) as the latent alignments, considering its flexibility of variable length prediction. CTC can efficiently find all valid aligned sequences at which the target $\boldsymbol{y}$ can be recovered and marginalize the log-likelihood.

2.3 Inference

Our decoding method during inference is simple and fast. The predicted vector $\boldsymbol{p} \in \mathbb{R}^{d_V}$ is converted to token ids by applying `argmax` function. The BERT special tokens such as `[CLS]` and `[PAD]` are removed from the generated texts. Following the previous methods (Gu and Kong, 2021), we removed repetitive tokens. The decoding of BioNART can be easily parallelized and has an advantage in generation speed over AR models.

2.4 Parameter Initialization

Loading Parameters from Publicly Available Checkpoints Each of the encoder and decoder is compatible with the widely used BERT, so we can initialize the parameters of BioNART with publicly available biomedical domain-specific PLMs such as BioBERT (Lee et al., 2019) and SciBERT (Beltagy et al., 2019). This parameter initialization is expected to boost the performance from a randomly initialized model and lower the cost of subsequent additional pre-training.

Additional Pre-training In order to familiarize our BioNART with NAR decoding, we performed an additional pre-training. We adopt a permutation language modeling strategy proposed in XLNet (Yang et al., 2019). BioNART is pre-trained on the PubMed/MEDLINE abstracts so that it can correctly rearrange the permutated tokens.

3 Experimental Settings

3.1 Loading PLMs Checkpoints

We initialized each of the encoder and decoder parts of BioNART with BioBERT (Lee et al., 2019) v1.1 base-cased checkpoints.

3.2 Our additional pre-training

We then performed the additional pre-training on the PubMed/MEDLINE abstract corpus[1]. This corpus contains 5.4B tokens of research article abstracts from the 2021 version of PubMed/MEDLINE. We used the same vocabulary as BioBERT (Lee et al., 2019) to tokenize the texts. We truncated all the input texts to 512 maximum sequence lengths following BioBART. We permuted the 30% of the input tokens, and the model is pre-trained with the CTC loss so that the tokens can be rearranged into the correct orders.

We set the mini-batch size to 2,560 by adopting gradient accumulation, which is the same size as BioBART. We used a learning rate scheduler with 0.1 warmup ratio and linear decay. The peak of the learning rate was 1e-04. We set the number of training epochs to 5, which corresponds to 45K steps. The number of training epochs is approximately 37.5% of the BioBART. We perform the additional pre-training on 16 40GB A100 GPUs for about 40 hours.

3.3 Datasets for Downstream Task

For fine-tuning on downstream tasks, we follow the same settings as BioBART. We fine-tuned our model on two biomedical language generation tasks in English: text summarization and clinical dialogue. For biomedical text summarization, we trained and evaluated BioNART on iCliniq and HealthCareMagic (Zeng et al., 2020), and for the clinical dialogue task, we employed CovidDialog (Ju et al., 2020) dataset. The details and statistics of the datasets are shown in Appendix A.1.

4 Results

4.1 Summarization

Table 1 shows the performance comparison between AR methods and our BioNART. First, our model is about eighteen times faster than AR models in inference. As measured with Green Algorithms (Lannelongue et al., 2021)[2], we can achieve a reduction of 0.13kg CO2 emission compared to BioBART model during the inference on HealthCareMagic test set. Regarding generative performance, BioNART showed the ROUGE-1 scores of about 81% and 78% of the performance by BioBART on the iCliniq and HealthCareMagic datasets, respectively. We present generative examples on iCliniq dataset in Appendix A.2 Table 3 shows the performance comparison of parameter initialization methods for the encoder and decoder parts of BioNART. These results show that loading the parameter checkpoints of BioBERT improved the ROUGE-score on iCliniq dataset.

By performing our additional pre-training, ROUGE-1/2/L scores improved 7.93/6.45/7.07 points and 3.38/2.62/2.88 points on the iCliniq and HealchCareMagic datasets, respectively. BioNART with additional pre-training model showed the ROUGE-1 scores of about 94% and 85% of the performance by BioBART on the iCliniq and Health-

[1] https://huggingface.co/datasets/pubmed

[2] http://calculator.green-algorithms.org

Model	iCliniq			
	ROUGE-1	ROUGE-2	ROUGE-L	latency
BART base (Lewis et al., 2020)	61.43	48.68	59.71	x18.5
BioBART base (Yuan et al., 2022)	61.07	48.47	59.42	x18.7
Mrini et al. (2021)	62.3	48.7	58.5	-
T5 base (Raffel et al., 2020)	58.28	46.09	55.28	x27.7
SciFive base (Phan et al., 2021)	55.63	43.13	52.07	x28.8
BioNART	49.50	37.54	48.17	**x1.0**
BioNART + our pre-training	57.43	43.99	55.24	**x1.0**
	HealthCareMagic			
	ROUGE-1	ROUGE-2	ROUGE-L	latency
BART base (Lewis et al., 2020)	46.81	26.19	44.34	x18.1
BioBART base (Yuan et al., 2022)	46.67	26.03	44.11	x18.9
Mrini et al. (2021)	46.9	24.8	43.2	-
BioNART	36.67	16.76	35.39	**x1.0**
BioNART + our pre-training	40.05	19.38	38.27	**x1.0**

Table 1: The main results on the summarization tasks.

Model	CovidDialogue				
	ROUGE-1	ROUGE-2	ROUGE-L	BLEU-4	latency
BART base (Lewis et al., 2020)	27.24	12.31	25.66	10.36	x19.1
BioBART base (Yuan et al., 2022)	28.14	12.77	26.32	11.40	x18.9
T5 base (Raffel et al., 2020)	20.14	8.75	16.68	6.62	x28.3
SciFive base (Phan et al., 2021)	22.29	9.60	18.65	7.13	x28.3
BioNART	2.91	0.00	2.91	0.00	**x1.0**
BioNART + our pre-training	15.19	1.21	10.58	0.75	**x1.0**

Table 2: The main results on the dialogue task.

Model	iCliniq		
	R-1	R-2	R-L
BioNART (random)	44.43	35.51	43.43
BioNART (BioBERT)	49.50	37.54	48.17

Table 3: The performance comparison between random initialization and BioBERT checkpoints initialization.

CareMagic datasets. These results show that our additional pre-training is effective for NAR generation tasks.

4.2 Dialogue

We show the performance comparison with AR models on the CovidDialogue dataset in Table 2. On the dataset, our model had a more significant performance difference from the AR model compared to the biomedical summarization datasets. We consider one of the factors is the small data size.

5 Related Work

Existing work has shown that pre-training the language models on the domain-specific texts can bring better performance on corresponding domain downstream tasks. BioBERT (Lee et al., 2019) is pre-trained on the PubMed abstracts and PubMed Central (PMC) full-text articles. Instead of continuous training from the general BERT checkpoint, SciBERT (Beltagy et al., 2019) and PubMedBERT (Gu et al., 2021) are trained from scratch from scientific papers on Semantic Scholar (Ammar et al., 2018) and PubMed papers, respectively.

BioBART (Yuan et al., 2022) further analyzed the influence of the pre-training task on the biomedical natural language generation (NLG) tasks and collated existing biomedical NLG tasks along with corresponding data and experimental settings, but all biomedical NLG models are AR-based, and there has been no discussion and verification of the NAR methods in the biomedical domain.

6 Conclusion

In this study, we proposed a non-autoregressive language model for the biomedical domain for the first time. Our method is based on the encoder-decoder model, and both the encoder and decoder are compatible with widely used BERT architecture, which allows for boosting generation performance and reducing additional pre-training costs. In the experiments on biomedical summarization tasks, BioNART achieved an 18 times inference speedup without a large performance drop compared to AR models. Our code and models are available at https://github.com/aistairc/BioNART

In future work, we would like to approach other downstream tasks such as relation extraction and entity linking as sequence-to-sequence problems and realize a high-speed and versatile model for the biomedical domain.

Limitations

This paper shows that NAR generation can be achieved from publicly available PLMs parameters. However, there is still room to be validated for additional pre-training. The performed pre-training in this study is small, and further investigation of whether the training steps and the size of the corpus are sufficient, and whether the self-training tasks other than permutation language modeling are effective.

Acknowledgements

This paper is based on results obtained from a project JPNP20006, commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

Waleed Ammar, Dirk Groeneveld, Chandra Bhagavatula, Iz Beltagy, Miles Crawford, Doug Downey, Jason Dunkelberger, Ahmed Elgohary, Sergey Feldman, Vu Ha, Rodney Kinney, Sebastian Kohlmeier, Kyle Lo, Tyler Murray, Hsu-Han Ooi, Matthew Peters, Joanna Power, Sam Skjonsberg, Lucy Lu Wang, Chris Wilhelm, Zheng Yuan, Madeleine van Zuylen, and Oren Etzioni. 2018. Construction of the literature graph in semantic scholar. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 3 (Industry Papers)*, pages 84–91, New Orleans - Louisiana. Association for Computational Linguistics.

Iz Beltagy, Kyle Lo, and Arman Cohan. 2019. SciBERT: A pretrained language model for scientific text. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3615–3620, Hong Kong, China. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Marjan Ghazvininejad, Omer Levy, Yinhan Liu, and Luke Zettlemoyer. 2019. Mask-predict: Parallel decoding of conditional masked language models. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 6112–6121, Hong Kong, China. Association for Computational Linguistics.

Alex Graves, Santiago Fernández, Faustino Gomez, and Jürgen Schmidhuber. 2006. Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, page 369–376, New York, NY, USA. Association for Computing Machinery.

Jiatao Gu and Xiang Kong. 2021. Fully non-autoregressive neural machine translation: Tricks of the trade. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 120–133, Online. Association for Computational Linguistics.

Yu Gu, Robert Tinn, Hao Cheng, Michael Lucas, Naoto Usuyama, Xiaodong Liu, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. 2021. Domain-specific language model pretraining for biomedical natural language processing. *ACM Transactions on Computing for Healthcare (HEALTH)*, 3(1):1–23.

Zeqian Ju, Subrato Chakravorty, Xuehai He, Shu Chen, Xingyi Yang, and Pengtao Xie. 2020. Covidialog: Medical dialogue datasets about covid-19. *https://github.com/UCSD-AI4H/COVID-Dialogue*.

Loïc Lannelongue, Jason Grealey, and Michael Inouye. 2021. Green algorithms: quantifying the carbon footprint of computation. *Advanced science*, 8(12):2100707.

Jason Lee, Elman Mansimov, and Kyunghyun Cho. 2018. Deterministic non-autoregressive neural sequence modeling by iterative refinement. In *Proceedings of the 2018 Conference on Empirical Methods*

in Natural Language Processing, pages 1173–1182, Brussels, Belgium. Association for Computational Linguistics.

Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang. 2019. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4):1234–1240.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.

Junyi Li, Tianyi Tang, Wayne Xin Zhao, Jian-Yun Nie, and Ji-Rong Wen. 2022. ELMER: A non-autoregressive pre-trained language model for efficient and effective text generation. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 1044–1058, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Himanshu Maheshwari, Bhavyajeet Singh, and Vasudeva Varma. 2021. SciBERT sentence representation for citation context classification. In *Proceedings of the Second Workshop on Scholarly Document Processing*, pages 130–133, Online. Association for Computational Linguistics.

Khalil Mrini, Franck Dernoncourt, Seunghyun Yoon, Trung Bui, Walter Chang, Emilia Farcas, and Ndapa Nakashole. 2021. A gradually soft multi-task and data-augmented approach to medical question understanding. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1505–1515, Online. Association for Computational Linguistics.

Long N. Phan, James T. Anibal, Hieu Tran, Shaurya Chanana, Erol Bahadroglu, Alec Peltekian, and Grégoire Altan-Bonnet. 2021. Scifive: a text-to-text transformer model for biomedical literature.

Weizhen Qi, Yeyun Gong, Jian Jiao, Yu Yan, Weizhu Chen, Dayiheng Liu, Kewen Tang, Houqiang Li, Jiusheng Chen, Ruofei Zhang, Ming Zhou, and Nan Duan. 2021. BANG: bridging autoregressive and non-autoregressive generation with large scale pre-training. In *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pages 8630–8639. PMLR.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, 21(1):5485–5551.

Sascha Rothe, Shashi Narayan, and Aliaksei Severyn. 2020. Leveraging pre-trained checkpoints for sequence generation tasks. *Transactions of the Association for Computational Linguistics*, 8:264–280.

Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R Salakhutdinov, and Quoc V Le. 2019. Xlnet: Generalized autoregressive pretraining for language understanding. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.

Hongyi Yuan, Zheng Yuan, Ruyi Gan, Jiaxing Zhang, Yutao Xie, and Sheng Yu. 2022. BioBART: Pretraining and evaluation of a biomedical generative language model. In *Proceedings of the 21st Workshop on Biomedical Language Processing*, pages 97–109, Dublin, Ireland. Association for Computational Linguistics.

Guangtao Zeng, Wenmian Yang, Zeqian Ju, Yue Yang, Sicheng Wang, Ruisi Zhang, Meng Zhou, Jiaqi Zeng, Xiangyu Dong, Ruoyu Zhang, Hongchao Fang, Penghui Zhu, Shu Chen, and Pengtao Xie. 2020. MedDialog: Large-scale medical dialogue datasets. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 9241–9250, Online. Association for Computational Linguistics.

A Appendix

In this section, we provide the details of the biomedical generation task datasets. The statistics of the datasets are shown in Table 4.

A.1 Dataset Details

A.1.1 Summarization

iCliniq and HealthCareMagic datasets are extracted from the MedDialog dataset (Zeng et al., 2020), collected from the online healthcare platform. iCliniq contains 31,062 samples, and HalthCareMagic contains 226,405 samples. Each sample comprises a summary and corresponding dialogues between a patient and a doctor. The summaries in HealthCareMagic are more abstractive, and they are written in a formal style, while the summaries in iCliniq are written by the patients. We follow previous work (Mrini et al., 2021; Yuan et al., 2022) for training, developing, and testing data separations of both datasets.

Dataset	Train	Dev	Test
iCliniq	24,851	3,105	3,108
HealthCareMagic	181,122	22,641	22,642
CovidDialog	490	63	61

Table 4: The statistics of the datasets for biomedical generative tasks.

A.1.2 Dialogue

CovidDialog (Ju et al., 2020) is a medical dialogue task that aims to imitate the human doctor to communicate with human patients in a natural way. The objective of this task is to correctly generate the replies by the doctors from the primitive descriptions of the patients and dialogue. The dataset is collected from online healthcare forums. Each consultation starts with a description related to the medical conditions of the patients, then follows the conversation between a doctor and a patient. The number of samples is 614, which is very small compared to the summarization task datasets.

A.2 Generated Examples

We show three randomly chosen examples from iCliniq dataset in Table 5. In the first example, both BioBART and BioNART generate the summary of asking about the causes of chest pain. However, the example generated by BioNART has a grammatical error. In the second example, both BioBART and BioNART summarized only the first part of the context and do not generate an adequate summary. In the third example, while BioBART model contains errors on the subject, BioNART model provides a more accurate summary.

Context	Gold summary	BioBART output	BioNART output
Patient: Hello doctor, I am suffering from chest pain in the left side from 4 to 5 days, but my lipid profile and 2D echocardiography both reports are normal. What to do? What is the reason of chest pain? Doctor: Hi. Chronic reflux of stomach acids into your esophagus can irritate your esophagus. In severe cases, it can cause dysphagia. Scar tissue can develop in your esophagus. The scar tissue can narrow your esophagus. This is known as an esophageal stricture. In some cases, dysphagia can be a direct result of esophageal damage. The lining of the esophagus can alter to resemble the tissue that lines your intestines. This is a condition called Barretts esophagus.	*I have chest pain despite lipid profile and echo being normal. Why?*	*What is the reason for chest pain in the left side?*	*What is the chest pain indicate?*
Patient: Hi doctor, My shoulder bone was dislocated before six months. I can move my hand upright, but while sleeping it is paining. What kind of deficiency can be there? I am eating non-vegetarian food regularly from childhood. Please suggest me some exercise and supplement as my bones are weak. Thank you in advance. Doctor: Hi. I have understood your concern. For further information consult a fitness expert online	*Please suggest some exercise and supplement for weak bones*	*I dislocated my shoulder bone. Please help*	*What is the specific deficiency for in hand in shoulder bone deficiency?*
Patient: My wife age 46. Suffering from heavy blooding during menstruation period. She came into same phase after eight months.(almost eight months no periods at all). Periods started on 13 days back, continue till date. Using 5-6 senetary pads during each day. Also passing through senior abdominal pain several days...also having complain of constipation since last three days. Consult local physician and than gynecologist in my city. She is swollowing medicines as per below... Tranexa mf/Trexakind thrice a day. Spasmodart/voveran if abdominal pain persists. Shelcal one per day. From yesterday she is taking injection of aquairon 25mg. (Three prescribed by local gynecologist. One taken) Also feeling weakness. I want second opinion and finest treatment suggestions online.... Please guide on above case and suggest finest treatment Doctor: Hello, Thanks for the query, With the history, it is clear that she is in her peri-menopausal status and experiencing menorrhagia or excessive bleeding since 13 days, which is often seen in perimenopausal women. She has been given antifibrinolytics in form or tranexa-mf which is fine, but if bleeding is still not getting controlled then she would require an ultrasound pelvis first to check for her present endometrial thickness . If the endometrial thickness is still more than 9 mm then she would have to be started on progesterone pills for 10 days to control bleeding along with diosmin and antifibrinolytics together. Also if she is feeling dizziness and weakness, then she may require hospitalization for intravenous support as well. Regards	*My wife is having heavy menstrual bleeding for 13 days. Please suggest treatment*	*I have heavy blooding during menstruation period. Please help*	*My wife suffering heavy blood bleeding during weakness. Please suggest*

Table 5: Example context-summary pairs and generated summaries from BioBART and BioNART on iCliniq development set.