Shimo Lab at “Discharge Me!”: Discharge Summarization by Prompt-Driven Concatenation of Electronic Health Record Sections

Yunzhen He[*1] **Hiroaki Yamagiwa**[*1] **Hidetoshi Shimodaira**[1,2]
[1] Kyoto University [2] RIKEN AIP
he.yunzhen.25d@st.kyoto-u.ac.jp,
hiroaki.yamagiwa@sys.i.kyoto-u.ac.jp, shimo@i.kyoto-u.ac.jp

Abstract

In this paper, we present our approach to the shared task “Discharge Me!” at the BioNLP Workshop 2024. The primary goal of this task is to reduce the time and effort clinicians spend on writing detailed notes in the electronic health record (EHR). Participants develop a pipeline to generate the “Brief Hospital Course” and “Discharge Instructions” sections from the EHR. Our approach involves a first step of extracting the relevant sections from the EHR. We then add explanatory prompts to these sections and concatenate them with separate tokens to create the input text. To train a text generation model, we perform LoRA fine-tuning on the ClinicalT5-large model. On the final test data, our approach achieved a ROUGE-1 score of 0.394, which is comparable to the top solutions.

1 Introduction

Electronic health records (EHR) eliminate the need for end-users to write medical records by hand and provide easy access to digital records (Menachemi and Collum, 2011). However, the use of EHR sometimes increases the burden on end-users (Shanafelt et al., 2016; Liu et al., 2022; Gao et al., 2023). With this in mind, there has been active research in recent years into applying natural language processing (NLP) to EHR to reduce the burden on end-users (Dong et al., 2022; Houssein et al., 2023; Veen et al., 2023).

To explore the potential of NLP in EHR, the shared task “Discharge Me!” (Xu et al., 2024) at the BioNLP Workshop 2024 evaluates the ability to generate discharge summaries. The goal of this task is to reduce the time and effort clinicians spend on writing detailed notes in the EHR. Participants develop a pipeline that leverages the EHR data to generate discharge summaries.

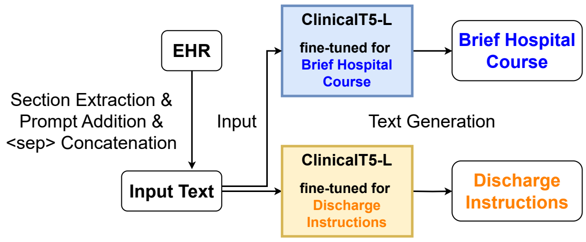


Figure 1: Overview of our pipeline. To create input text, we extract sections from the EHR, add explanatory prompts, and then concatenate them with <sep> tokens. We then generate discharge summaries using ClnicalT5-large, which has been fine-tuned for each target.

In this paper, we present our approach to the shared task. Fig. 1 provides an overview of our pipeline. We preprocess the EHR, as illustrated in Fig. 2, by removing noise and extracting sections that are essential for the target summary. The sections are selected based on a predetermined priority. For extracted sections, we prepend the prompt from Table 2 to the beginning of the text, concatenate these sections using <sep> tokens, and thus prepare the input text. We also removed noise from the target text. We then fine-tuned ClinicalT5 (Lu et al., 2022), which is pre-trained on clinical texts. On the final test data, our approach achieved a ROUGE-1 score of 0.394, which is comparable to the top solutions.

2 Related work

2.1 Text generation models in clinical domain

Decoder. ClinicalGPT (Wang et al., 2023), whose base model is BLOOM-7B (Le Scao et al., 2022), uses LoRA (Hu et al., 2022) for fine-tuning and applies the reinforcement learning process used in InstructGPT (Ouyang et al., 2022). BioMistral-7B (Labrak et al., 2024) underwent additional pretraining of the Mistral-7B (Jiang et al., 2023) model on PubMed Central (Roberts, 2001) and showed good performance on the clinical knowledge QA

[*] The first two authors contributed equally to this work. Our code is available at https://github.com/githubhyz/DischargeMe_BioNLP2024.

Proceedings of the 23rd Workshop on Biomedical Language Processing, pages 645–657
August 16, 2024.

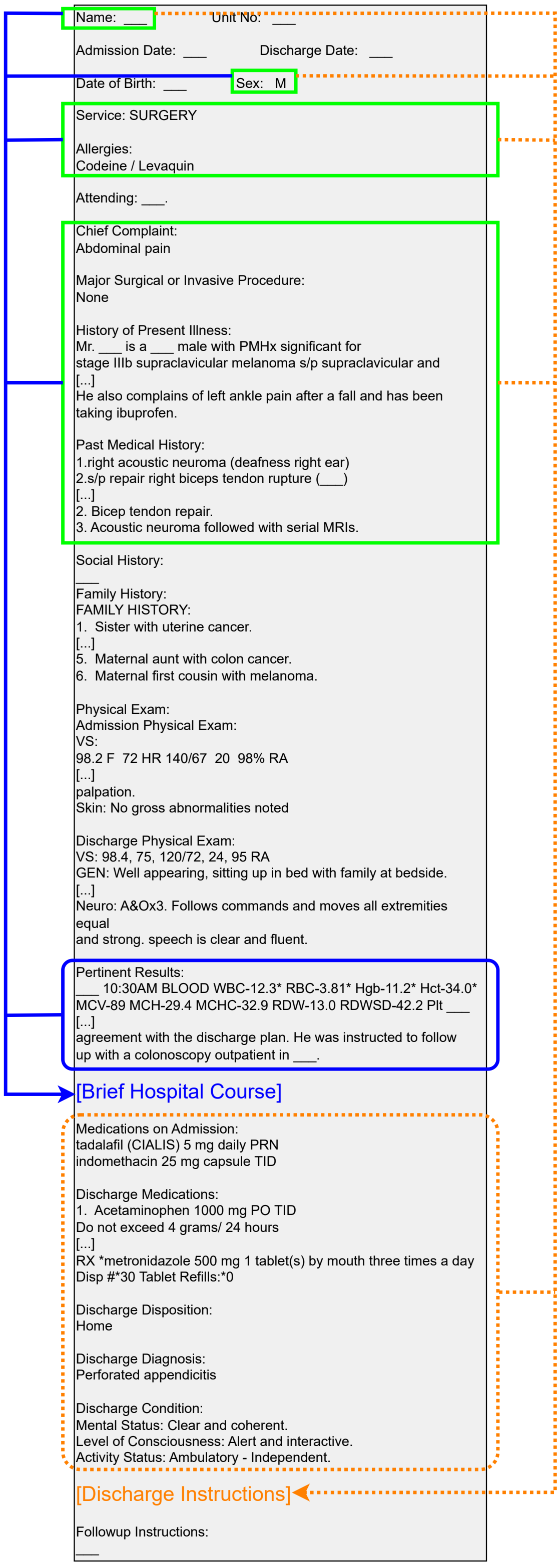


Figure 2: An example of the EHR with the location of the target discharge summaries. To show the sections used for the input text, the rounded rectangle is for the "Brief Hospital Course", the dashed rounded rectangle is for the "Discharge Instructions", and the rectangles are for both targets. The symbol "[...]" indicates omissions.

task.

Encoder-decoder. ClinicalT5 (Lu et al., 2022; Lehman et al., 2023), whose base model is T5 (Raffel et al., 2020), is the model pre-trained on clinical texts[1]. Lu et al. (2022) performed additional pre-training of the SciFive-PubMed-PMC (Phan et al., 2021) model on MIMIC-III (Johnson et al., 2016). Meanwhile, Lehman et al. (2023) pre-trained T5 from scratch using MIMIC-III and MIMIC-IV (Johnson et al., 2023c).

2.2 Clinical text summarization

Discharge Summarization. Williams et al. (2024) showed that although 33% of the discharge summaries generated by GPT-4 (Achiam et al., 2023) from the EHR were error-free, some contained hallucinations and omitted relevant information. Note, however, that the shared task does not allow data to be sent to third parties via an API.

Problem List Summarization (ProbSum). ProbSum (Gao et al., 2022) is a task aimed at generating a list of problems in a patient's daily care plan based on hospital records. In the BioNLP 2023 shared task (Gao et al., 2023) focused on ProbSum, the ensemble of ClinicalT5 models demonstrated robust performance (Manakul et al., 2023), and the approach combining Flan-T5 (Chung et al., 2022) with GPT2XL (Radford et al., 2019) also yielded strong results (Li et al., 2023). In the experiments using the shared task dataset, LLMs adapted to the medical domain demonstrated performance equal to or better than medical experts (Van Veen et al., 2024).

3 Task overview

3.1 Task description

Participants use an EHR dataset from MIMIC-IV (Johnson et al., 2023c) and develop a pipeline to generate two discharge summaries: the "Brief Hospital Course" section for patients and the "Discharge Instructions" section for clinicians. Table 1 shows an example of both sections.

3.2 Dataset description

The original datasets (Xu, 2024) include training, validation, phase I test, and phase II test sets. Participants use the training and validation sets to develop their pipeline, with the final evaluation per-

[1]Both of Lu et al. (2022) and Lehman et al. (2023) refer to their models as ClinicalT5.

Brief Hospital Course	Discharge Instructions
Mr. ___ is a ___ yo M with medical history significant for \\ stage IIIb supraclavicular melanoma and prostate cancer admitted \\ to the Acute Care Surgery Service on ___ with worsening \\ abdominal pain, frequent stools, and subjective fevers. He was \\ transferred from ___ for further management with a CT \\ abdomen showing a 5 x 6 x 7 cm right mid abdominal inflammatory \\ phlegmon. He was admitted to the surgical floor for IV \\ antibitoics and further evaluation.\\ \\ Gastroenterology was consulted for duodenal thickening. Given \\ his current infection the wall thickening is likely secondary to \\ the infection. Repeat imaging was recommended to evaluate \\ evolution of the phlegmon as well as outpatient colonoscopy once \\ antibiotic treatment is complete. \\ \\ The remainder of the hospital course is summarized below:\\ Neuro: The patient was alert and oriented throughout \\ hospitalization; pain was initially managed with a IV dilaudid. \\ He had left ankle pain and swelling consistent with gout that \\ was managed with PO indomethacin.. \\ CV: The patient remained stable from a cardiovascular \\ standpoint; vital signs were routinely monitored.\\ Pulmonary: The patient remained stable from a pulmonary \\ standpoint. Good pulmonary toilet, early ambulation and \\ incentive spirometry were encouraged throughout hospitalization. \\ \\ GI/GU/FEN: The patient was initially kept NPO. On HD3 he was \\ given a clear liquid diet. On HD4 he was advanced to regular \\ diet with good tolerability. Patient's intake and output were \\ closely monitored\\ ID: The patient's fever curves were closely watched for signs of \\ infection, of which there were none. He was initially given IV \\ zosyn and transitioned to oral flagyl and ciprofloxacin upon \\ discharge to complete a 2 week course of antibiotics. \\ HEME: The patient's blood counts were closely watched for signs \\ of bleeding, of which there were none.\\ Prophylaxis: The patient received subcutaneous heparin and ___ \\ dyne boots were used during this stay and was encouraged to get \\ up and ambulate as early as possible.\\ \\ At the time of discharge, the patient was doing well, afebrile \\ and hemodynamically stable. The patient was tolerating a diet, \\ ambulating, voiding without assistance, and pain was well \\ controlled. The patient received discharge teaching and \\ follow-up instructions with understanding verbalized and \\ agreement with the discharge plan. He was instructed to follow \\ up with a colonoscopy outpatient in ___.	Dr. ___,\\ \\ You were admitted to the Acute Care Surgery Service on ___ \\ with abdominal pain. You had a CT scan of your abdomen that \\ showed likely a perforated appendicitis. You were given IV \\ antibiotics and had improvement in your symptoms. An attempt was \\ made to drain the infection but it is not amenable to a drain at \\ this time. You were transitioned to oral antibiotics with \\ continued good effect.\\ \\ While in the hospital you had a flair up of gout in your left \\ ankle. You were given indomethacin with improvement in your \\ symptoms.\\ \\ You are now doing better, tolerating a regular diet, and ready \\ to be discharged to home to continue your recovery.\\ \\ Please note the following discharge instructions:\\ \\ Please call your doctor or nurse practitioner or return to the \\ Emergency Department for any of the following:\\ *You experience new chest pain, pressure, squeezing or \\ tightness.\\ *New or worsening cough, shortness of breath, or wheeze.\\ *If you are vomiting and cannot keep down fluids or your \\ medications.\\ *You are getting dehydrated due to continued vomiting, diarrhea, \\ or other reasons. Signs of dehydration include dry mouth, rapid \\ heartbeat, or feeling dizzy or faint when standing.\\ *You see blood or dark/black material when you vomit or have a \\ bowel movement.\\ *You experience burning when you urinate, have blood in your \\ urine, or experience a discharge.\\ *Your pain in not improving within ___ hours or is not gone \\ within 24 hours. Call or return immediately if your pain is \\ getting worse or changes location or moving to your chest or \\ back.\\ *You have shaking chills, or fever greater than 101.5 degrees \\ Fahrenheit or 38 degrees Celsius.\\ *Any change in your symptoms, or any new symptoms that concern \\ you.\\ \\ Please resume all regular home medications, unless specifically \\ advised not to take a particular medication. Also, please take \\ any new medications as prescribed.\\ \\ Please get plenty of rest, continue to ambulate several times \\ per day, and drink adequate amounts of fluids.

Table 1: An example of the "Brief Hospital Course" and "Discharge Instructions" sections. "\\" means line breaks.

Section	Prompt	Brief Hospital Course	Discharge Instructions
Name	The patient's name is provided as follows:	1	1
Sex	Gender details are as follows:	2	2
Service	The service details are as follows:	9	9
Allergies	Information on any allergies is detailed as follows:	7	6
Chief Complaint	The primary reason for the visit is summarized as follows:	3	3
Major Surgical or Invasive Procedure	Details on any major surgeries or invasive procedures are as follows:	8	7
History of Present Illness	An overview of the current illness's history is provided as follows:	4	4
Past Medical History	A summary of the patient's past medical history is as follows:	6	5
Pertinent Results	Clinically significant findings impacting the treatment and diagnosis are as follows:	5	–
Medications on Admission	Medications upon admission are detailed as follows:	–	8
Discharge Diagnosis	The final diagnosis at discharge is as follows:	–	10
Discharge Disposition	The disposition at discharge is provided as follows:	–	11
Discharge Condition	The patient's condition upon discharge is described as follows:	–	12
Discharge Medications	Medications prescribed at discharge are as follows:	–	13

Table 2: Prompts for each section and their priorities in each target discharge summary. The priority is used to order the sections in the input text.

formed on a subset of 250 samples from the phase II test set. See Appendix A for more details.

Note that although the datasets include metadata such as radiology reports in addition to the EHR and discharge summaries, we did not use this information in designing a simple pipeline. For more details, see the task website[2].

[2] https://stanford-aimi.github.io/discharge-me/

We created a new split with a 4:1 training-to-validation ratio using the original training and validation sets. Note that the EHR in the dataset contains the target texts: the "Brief Hospital Course" and the "Discharge Instructions" sections. As shown in Fig. 2, the "Brief Hospital Course" section is usually located in the middle of the discharge summary, while the "Discharge Instructions" section is generally located at the end of the EHR.

Rank	Team	Overall	BLEU	ROUGE-1	ROUGE-2	ROUGE-L	BERTScore	Meteor	AlignScore	MEDCON
1	WisPerMed	**0.332**	**0.124**	**0.453**	**0.201**	**0.308**	**0.438**	**0.403**	**0.315**	**0.411**
2	HarmonAI Lab at Yale	0.300	0.106*	0.423	0.180*	0.284	0.412	0.381	0.265*	0.353*
3	aehrc	0.297*	0.097	0.414	0.192	0.284	0.383*	0.398	0.274*	0.332*
4	EPFL-MAKE	0.289*	0.098	0.444	0.155	0.262*	0.399	0.336*	0.255*	0.360*
5	UF-HOBI	0.286*	0.102*	0.401*	0.174*	0.275*	0.395	0.289	0.296	0.355*
6	de ehren	0.284*	0.097	0.404*	0.166*	0.265*	0.389*	0.376	0.231	0.339*
7	DCT_PI	0.277*	0.092	0.401*	0.158	0.256*	0.378*	0.363	0.247	0.320
8	IgnitionInnovators	0.253	0.068	0.370*	0.131	0.245	0.360*	0.314	0.215	0.324
9	Shimo Lab (Ours)	0.248	0.063	0.394*	0.131	0.252*	0.351*	0.312	0.210	0.276
10	qub-cirdan	0.221	0.024	0.377*	0.106	0.205	0.300	0.332*	0.174	0.254
11	Roux-lette	0.206	0.030	0.319	0.084	0.182	0.289	0.287	0.195	0.265
12	UoG Siephers	0.191	0.017	0.341	0.109	0.209	0.268	0.247	0.143	0.193
13	mike-team	0.188	0.022	0.290	0.076	0.163	0.258	0.294	0.182	0.223
14	Ixa-UPV	0.183	0.016	0.259	0.057	0.144	0.282	0.284	0.210	0.215
15	MLBMIKABR	0.170	0.039	0.210	0.092	0.131	0.186	0.306	0.205	0.191
16	cyq	0.104	0.002	0.197	0.016	0.106	0.179	0.106	0.132	0.091
17	AIMI-Baseline	0.102	0.015	0.126	0.052	0.113	0.138	0.089	0.167	0.121

Table 3: The evaluation metrics values for the final test data. The higher values are better, and the highest value is highlighted in **bold**. Values that are at least 90% of the highest value are underlined, and values that are at least 80% of the highest value are marked with $(*)$.

Brief Hospital Course	Discharge Instructions
Mr. ___ is a __ yo M with PMHx significant for stage IIIb supraclavicular melanoma s/p supraclavicle and right anterior neck dissection and prostate cancer presenting with abdominal pain. The remainder of the hospital course is summarized below: - Neuro: The patient was alert and oriented throughout hospitalization; pain was initially managed with a IV dilaudid. He had left ankle pain and swelling consistent with gout that was managed with PO indomethacin. CV: The patient remained stable from a cardiovascular standpoint; vital signs were routinely monitored. Pulmonary: - The patient stayed stable from an pulmonary standpoint. Good pulmonary toilet, early ambulation and incentive spirometry were encouraged throughout hospitalization - GI/GU/FEN: The patient is initially kept NPO. On HD3 he was given a clear liquid diet. On HD4 he was advanced to regular diet with good tolerability. - Patient's intake and output were closely monitored ID: The patient's fever curves were closely watched for signs of infection, of which there were none. He was initially given IV zosyn and transitioned to oral flagyl and ciprofloxacin upon discharge to complete a 2 week course of antibiotics. HEME: The patient received subcutaneous heparin and dyne boots were used during this stay and was encouraged to get up and ambulate as early as possible. At the time of discharge, the patient was doing well, afebrile and hemodynamically stable. The patient was tolerating a diet, ambulating, voiding without assistance, and pain was well controlled. The patient received discharge teaching and follow-up instructions with understanding verbalized and agreement with the discharge plan. He was instructed to follow up with a colonoscopy outpatient in ___	Dear Mr. ___, You were admitted to the hospital with abdominal pain. You were found to have a perforated appendicitis. You were treated with bowel rest and intravenous antibiotics. You are now ready to be discharged home to continue your recovery with the following instructions: Please call your doctor or nurse practitioner or return to the Emergency Department for any of the following: *You experience new chest pain, pressure, squeezing or tightness. *New or worsening cough, shortness of breath, or wheeze. *If you are vomiting and cannot keep down fluids or your medications. *You are getting dehydrated due to continued vomiting, diarrhea, or other reasons. Signs of dehydration include dry mouth, rapid heartbeat, or feeling dizzy or faint when standing. *You see blood or dark/black material when you vomit or have a bowel movement. *You experience burning when you urinate, have blood in your urine, or experience a discharge. *Your pain in not improving within 12 hours or is not gone within 24 hours. Call or return immediately if your pain is getting worse or changes location or moving to your chest or back. *You have shaking chills, or fever greater than 101.5 degrees Fahrenheit or 38 degrees Celsius. *Any change in your symptoms, or any new symptoms that concern you. Please resume all regular home medications, unless specifically advised not to take a particular medication. Also, please take any new medications as prescribed. Please get plenty of rest, continue to ambulate several times per day, and drink adequate amounts of fluids. Avoid lifting weights greater than __- lbs until you follow-up with your surgeon. Avoid driving or operating heavy machinery while taking pain medications.

Table 4: Our generated texts for the "Brief Hospital Course" and "Discharge Instructions" sections in Table 1.

3.3 Evaluation metrics

In this task, the following eight evaluation metrics[3] are used to compare the generated texts with the target texts: BLEU-4 (Papineni et al., 2002), ROUGE-1, ROUGE-2, ROUGE-L (Lin, 2004), BERTScore (Zhang et al., 2020), METEOR (Banerjee and Lavie, 2005), AlignScore (Zha et al., 2023), MEDCON (Yim et al., 2023). The overall score is calculated by first averaging the scores for each target, and then averaging these values.

4 Pipeline

4.1 Input text preparation

We removed the target discharge summaries from the EHR as preprocessing. As shown in Fig. 2, the EHR contains redundant line breaks and detailed data. When the EHR is used directly as input text, this redundancy can increase the length of the input text. To mitigate this, we removed the noise from the EHR and selectively extracted the relevant sections for each target, thus avoiding the excessive length of the input text[4]. These sections were selected by excluding those with detailed data, such as timestamps[5], or those without specific information, such as the "Admission Date" section. Note that, in the case of preparing the input text for the model generating the "Brief Hospital Course" section, given the actual workflow of writing discharge summaries, we did not use the sections following this section in the input text.

For sections extracted from the EHR, we added an explanatory prompt to the beginning of each section and then concatenated the sections using the

[3] `https://github.com/Stanford-AIMI/discharge-me/tree/main/scoring`.

[4] The criteria for section selection are ad hoc, as mentioned in the Limitations section.

[5] Although the "Pertinent Results" section contains timestamps, we exclude them and use this section as input for the "Brief Hospital Course" section. See the Appendix B.3 for details.

<sep> tokens to create the final input text. Table 2 shows the prompts and priorities of the selected sections used in the input text for each target discharge summary. The sections in the input text were ordered according to the specified priorities, rather than their original order in the EHR. The input text was truncated if it exceeded the maximum text length[6].

In Appendix B, examples of input texts are shown in Tables 6 and 7, respectively, for "Brief Hospital Course" and "Discharge Instructions". These input texts were prepared from the EHR in Fig. 2. Histograms of the length of the input text are shown in Fig. 3.

4.2 Target text preparation

As shown in Table 1, the target texts contain many unnecessary line breaks. To prevent the line breaks from hindering the learning of the model, we removed them during preprocessing. In Appendix C, the texts before and after preprocessing for "Brief Hospital Course" are shown in Table 9, and those for "Discharge Instructions" are shown in Table 10. Histograms of the length of the target text are shown in Fig. 4.

4.3 Text generation

Using the input and target texts prepared in Sections 4.1 and 4.2, we performed LoRA (Hu et al., 2022) fine-tuning on the ClinicalT5-large[7] model published by Lu et al. (2022). The ClinicalT5-large model has 770M parameters with 24 layers. In Appendix D, the hyperparameters for fine-tuning and LoRA are shown in Tables 13 and 14. The hyperparameters to generate each target discharge summary are shown in Table 15.

5 Experiments

5.1 Results for the final test data

Table 3 presents the evaluation metrics values of the participating teams for the final test data. While our method did not achieve the highest scores of *WisPerMed* (Damm et al., 2024), it demonstrated relatively good performance in ROUGE-1, ROUGE-L, and BERTScore. In particular, we achieved a ROUGE-1 score of 0.394, which is comparable to top solutions such as those of *HarmonAI Lab at Yale* and *aehrc*.

[6]1596 tokens
[7]https://huggingface.co/luqh/ClinicalT5-large

5.2 Qualitative observation

Table 4 presents the summaries generated by our pipeline from the EHR for the target summaries in Table 1. While the detailed progress reports and discharge instructions may differ, the overall gist remains the same. In addition, unnecessary line breaks that were present in the original target summaries do not appear in the generated summaries.

6 Conclusion

We presented our approach to the shared task "Discharge Me!" at the BioNLP Workshop 2024. Extracting the relevant sections from the EHR, we added explanatory prompts to these sections and concatenated them with <sep> tokens to create the input text. We then performed LoRA fine-tuning on the ClinicalT5-large model. On the final test data, our approach achieved a ROUGE-1 score of 0.394, which is comparable to the top solutions.

Limitations

- Our pipeline cannot be applied to an EHR with different formats, resulting in a lack of generalizability. In fact, even in this shared task dataset, the lack of consistency in the original data sometimes makes it impossible to extract sections, resulting in incomplete summaries.
- When preparing the input text, adding prompts for each extracted section results in a longer length than simply concatenating sections with <sep> tokens.
- The effectiveness of our pipeline is not tested against other text generation models such as BioMistral-7B (Labrak et al., 2024) and the ClinicalT5-large model published by Lehman et al. (2023).
- While the selection and prioritization of the EHR sections used in the input text is somewhat ad-hoc, since extensive experiments would be required to compare the selection and prioritization, we did not conduct them in this study due to time and resource constraints.
- While the cleaned target texts are used for training, the original target texts with many line breaks are used for evaluation. This leads to a discrepancy between the target text distributions of training and evaluation.

Ethics Statement

We conducted our research with careful consideration of data use and in accordance with the Data Use Agreement[8]. It is prohibited to identify individuals or organizations from the examples presented in the paper.

Acknowledgements

We would like to thank the Program Committee for their thorough review and valuable suggestions. This work was supported by JST SPRING, Grant Number JPMJSP2110. This study was partially supported by JST CREST JPMJCR21N3.

References

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Satanjeev Banerjee and Alon Lavie. 2005. METEOR: an automatic metric for MT evaluation with improved correlation with human judgments. In *Proceedings of the Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization@ACL 2005, Ann Arbor, Michigan, USA, June 29, 2005*, pages 65–72. Association for Computational Linguistics.

Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Eric Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Y. Zhao, Yanping Huang, Andrew M. Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. 2022. Scaling instruction-finetuned language models. *CoRR*, abs/2210.11416.

Hendrik Damm, Tabea Margareta Grace Pakull, Bahadir Eryilmaz, Helmut Becker, Ahmad Idrissi-Yaghir, Henning Schäfer, Sergej Schultenkämper, and Christoph M. Friedrich. 2024. Wispermed at "discharge me!": Advancing text generation in healthcare with large language models, dynamic expert selection, and priming techniques on MIMIC-IV. *CoRR*, abs/2405.11255.

Hang Dong, Matús Falis, William Whiteley, Beatrice Alex, Joshua Matterson, Shaoxiong Ji, Jiaoyan Chen, and Honghan Wu. 2022. Automated clinical coding: what, why, and where we are? *npj Digit. Medicine*, 5.

[8]https://physionet.org/content/discharge-me/view-dua/1.3/

Yanjun Gao, Dmitriy Dligach, Timothy Miller, and Majid Afshar. 2023. Overview of the problem list summarization (probsum) 2023 shared task on summarizing patients' active diagnoses and problems from electronic health record progress notes. In *The 22nd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks, BioNLP@ACL 2023, Toronto, Canada, 13 July 2023*, pages 461–467. Association for Computational Linguistics.

Yanjun Gao, Dmitriy Dligach, Timothy Miller, Dongfang Xu, Matthew M. Churpek, and Majid Afshar. 2022. Summarizing patients' problems from hospital progress notes using pre-trained sequence-to-sequence models. In *Proceedings of the 29th International Conference on Computational Linguistics, COLING 2022, Gyeongju, Republic of Korea, October 12-17, 2022*, pages 2979–2991. International Committee on Computational Linguistics.

Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. 2000. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23):E215–E220.

Essam H. Houssein, Rehab E. Mohamed, and Abdelmgeid A. Ali. 2023. Heart disease risk factors detection from electronic health records using advanced nlp and deep learning techniques. *Scientific Reports*, 13(1):7173.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. Lora: Low-rank adaptation of large language models. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.

Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de Las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Lélio Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. 2023. Mistral 7b. *CoRR*, abs/2310.06825.

Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Leo Anthony Celi, Roger Mark, and Steven Horng. 2023a. MIMIC-IV-ED (version 2.2).

Alistair Johnson, Tom Pollard, Steven Horng, Leo Anthony Celi, and Roger Mark. 2023b. MIMIC-IV-Note: Deidentified free-text clinical notes (version 2.2).

Alistair E. W. Johnson, Lucas Bulgarelli, Lu Shen, Alvin Gayles, Ayad Shammout, Steven Horng, Tom J. Pollard, Sicheng Hao, Benjamin Moody, Brian Gow, Li-wei H. Lehman, Leo A. Celi, and Roger G. Mark. 2023c. Mimic-iv, a freely accessible electronic health record dataset. *Sci Data*, 10(1):1. Erratum in: Sci

Data. 2023 Jan 16;10(1):31. Erratum in: Sci Data. 2023 Apr 18;10(1):219.

Alistair E. W. Johnson, Tom J. Pollard, Lu Shen, Li-wei H. Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G. Mark. 2016. Mimic-iii, a freely accessible critical care database. *Sci Data*, 3:160035.

Yanis Labrak, Adrien Bazoge, Emmanuel Morin, Pierre-Antoine Gourraud, Mickael Rouvier, and Richard Dufour. 2024. Biomistral: A collection of open-source pretrained large language models for medical domains. *CoRR*, abs/2402.10373.

Teven Le Scao, Angela Fan, Christopher Akiki, Ellie Pavlick, Suzana Ilić, Daniel Hesslow, Roman Castagné, Alexandra Sasha Luccioni, François Yvon, Matthias Gallé, et al. 2022. Bloom: A 176b-parameter open-access multilingual language model.

Eric Lehman, Evan Hernandez, Diwakar Mahajan, Jonas Wulff, Micah J. Smith, Zachary Ziegler, Daniel Nadler, Peter Szolovits, Alistair E. W. Johnson, and Emily Alsentzer. 2023. Do we still need clinical language models? In *Conference on Health, Inference, and Learning, CHIL 2023, Broad Institute of MIT and Harvard (Merkin Building), 415 Main Street, Cambridge, MA, USA*, volume 209 of *Proceedings of Machine Learning Research*, pages 578–597. PMLR.

Hao Li, Yu-Ping Wu, Viktor Schlegel, Riza Batista-Navarro, Thanh-Tung Nguyen, Abhinav Ramesh Kashyap, Xiao-Jun Zeng, Daniel Beck, Stefan Winkler, and Goran Nenadic. 2023. Team: PULSAR at probsum 2023: PULSAR: pre-training with extracted healthcare terms for summarising patients' problems and data augmentation with black-box large language models. In *The 22nd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks, BioNLP@ACL 2023, Toronto, Canada, 13 July 2023*, pages 503–509. Association for Computational Linguistics.

Chin-Yew Lin. 2004. ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain. Association for Computational Linguistics.

Jinghui Liu, Daniel Capurro, Anthony Nguyen, and Karin Verspoor. 2022. "note bloat" impacts deep learning-based nlp models for clinical prediction tasks. *Journal of biomedical informatics*, 133:104149.

Qiuhao Lu, Dejing Dou, and Thien Huu Nguyen. 2022. Clinicalt5: A generative language model for clinical text. In *Findings of the Association for Computational Linguistics: EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 5436–5443. Association for Computational Linguistics.

Potsawee Manakul, Yassir Fathullah, Adian Liusie, Vyas Raina, Vatsal Raina, and Mark J. F. Gales. 2023. CUED at probsum 2023: Hierarchical ensemble of summarization models. In *The 22nd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks, BioNLP@ACL 2023, Toronto, Canada, 13 July 2023*, pages 516–523. Association for Computational Linguistics.

Sourab Mangrulkar, Sylvain Gugger, Lysandre Debut, Younes Belkada, Sayak Paul, and Benjamin Bossan. 2022. Peft: State-of-the-art parameter-efficient fine-tuning methods. `https://github.com/huggingface/peft`.

Nir Menachemi and Taleah H Collum. 2011. Benefits and drawbacks of electronic health record systems. *Risk Manag Healthc Policy*, 4:47–55.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F. Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Z. Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. 2019. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pages 8024–8035.

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. *CoRR*, abs/2106.03598.

Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners.

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. *J. Mach. Learn. Res.*, 21:140:1–140:67.

Richard J Roberts. 2001. Pubmed central: The genbank of the published literature.

Tait D Shanafelt, Lotte N Dyrbye, Christine Sinsky, Omar Hasan, Daniel Satele, Jeff Sloan, and Colin P West. 2016. Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. *Mayo Clin Proc*, 91(7):836–848.

Dave Van Veen, Cara Van Uden, Louis Blankemeier, Jean-Benoit Delbrouck, Asad Aali, Christian Bluethgen, Anuj Pareek, Malgorzata Polacin, Eduardo Pontes Reis, Anna Seehofnerová, et al. 2024. Adapted large language models can outperform medical experts in clinical text summarization. *Nature Medicine*, pages 1–9.

Dave Van Veen, Cara Van Uden, Louis Blankemeier, Jean-Benoit Delbrouck, Asad Aali, Christian Bluethgen, Anuj Pareek, Malgorzata Polacin, Eduardo Pontes Reis, Anna Seehofnerova, Nidhi Rohatgi, Poonam Hosamani, William Collins, Neera Ahuja, Curtis P. Langlotz, Jason Hom, Sergios Gatidis, John Pauly, and Akshay S. Chaudhari. 2023. Clinical text summarization: Adapting large language models can outperform human experts. *Res Sq*. Preprint: rs.3.rs-3483777.

Guangyu Wang, Guoxing Yang, Zongxin Du, Longjun Fan, and Xiaohu Li. 2023. Clinicalgpt: Large language models finetuned with diverse medical data and comprehensive evaluation. *CoRR*, abs/2306.09968.

Christopher YK Williams, Jaskaran Bains, Tianyu Tang, Kishan Patel, Alexa N Lucas, Fiona Chen, Brenda Y Miao, Atul J Butte, and Aaron E Kornblith. 2024. Evaluating large language models for drafting emergency department discharge summaries. *medRxiv*, pages 2024–04.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, EMNLP 2020 - Demos, Online, November 16-20, 2020*, pages 38–45. Association for Computational Linguistics.

Justin Xu. 2024. Discharge me: Bionlp acl'24 shared task on streamlining discharge documentation. `https://doi.org/10.13026/4a0k-4360`. Version 1.2.

Dataset	Numbers
Training	68,785
Validation	14,719
Phase I Test	14,702
Phase II Test	10,962
Total	109,168

Table 5: The number of samples for the data splits.

Justin Xu, Zhihong Chen, Andrew Johnston, Louis Blankemeier, Maya Varma, Jason Hom, William J. Collins, Ankit Modi, Robert Lloyd, Benjamin Hopkins, Curtis Langlotz, and Jean-Benoit Delbrouck. 2024. Overview of the first shared task on clinical text generation: Rrg24 and "discharge me!". In *The 23rd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks*, Bangkok, Thailand. Association for Computational Linguistics.

Wen-wai Yim, Yujuan Fu, Asma Ben Abacha, Neal Snider, Thomas Lin, and Meliha Yetisgen. 2023. Aci-bench: a novel ambient clinical intelligence dataset for benchmarking automatic visit note generation. *Scientific Data*, 10(1):586.

Yuheng Zha, Yichi Yang, Ruichen Li, and Zhiting Hu. 2023. Alignscore: Evaluating factual consistency with A unified alignment function. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 11328–11348. Association for Computational Linguistics.

Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. 2020. Bertscore: Evaluating text generation with BERT. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net.

A Details of datasets

In this task, we use the dataset created by the MIMIC-IV's submodules MIMIC-IV-ED (Johnson et al., 2023a) and MIMIC-IV-Note (Johnson et al., 2023b). The dataset is available on PhysioNet (Goldberger et al., 2000), and its use requires completion of the CITI[9] training and credentialing process. Table 5 lists the number of samples for the data splits.

B Details of input text

This section first explains the detailed preprocessing required to create input text from the EHR. It

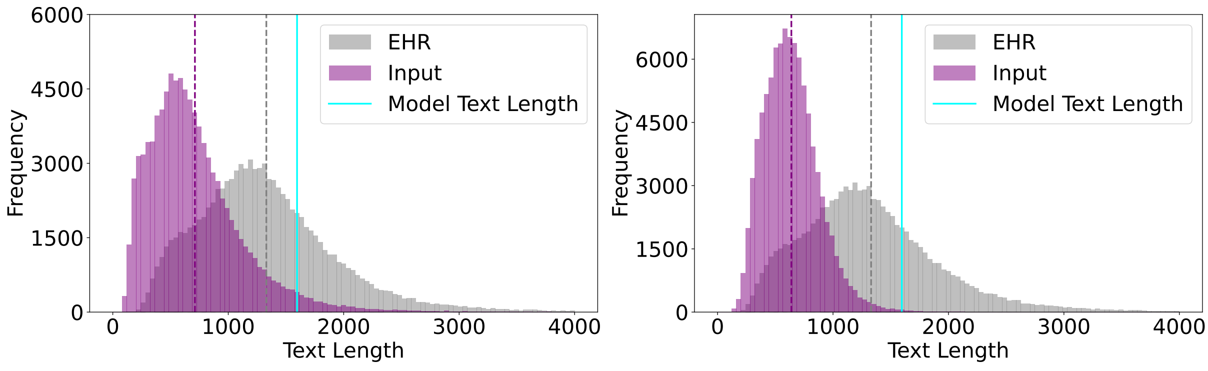
[9] `https://about.citiprogram.org/`

The patient's name is provided as follows: ___.<sep>Gender details are as follows: Male.<sep>The primary reason for the visit is summarized as follows: Abdominal pain.<sep>An overview of the current illness's history is provided as follows: Mr. ___ is a ___ male with PMHx significant for stage IIIb supraclavicular melanoma s/p supraclavicular and right anterior neck dissection and prostate cancer s/p radical prostatectomy, presenting with abdominal pain. The pain began two weeks ago and has been worsening in the last few days. His pain is localized to the right periumbilical region. He endorses having chills, inability to sleep and eat due to pain and a 6 lb weight loss in the past few days. He has been passing flatus and having loose and frequent non-bloody stools. He has also been having night sweats. He denies fever, nausea or vomiting. He was seen at ___ and transferred to ___ ED for further management after his CT abdomen showed a 5 x 6 x 7 cm right mid abdominal inflammatory phlegmon. His last colonoscopy was ___ years ago which revealed some polyps. He also complains of left ankle pain after a fall and has been taking ibuprofen.<sep>Clinically significant findings impacting the treatment and diagnosis are as follows: MCV-89 MCH-29.4 MCHC-32.9 RDW-13.0 RDWSD-42.2 Plt ___ MCV-88 MCH-29.8 MCHC-34.1 RDW-12.7 RDWSD-40.3 Plt ___ MCV-88 MCH-29.9 MCHC-33.8 RDW-12.5 RDWSD-40.2 Plt ___ MCV-88 MCH-29.5 MCHC-33.5 RDW-12.4 RDWSD-39.8 Plt ___ K-3.8 Cl-103 HCO3-24 AnGap-18 K-3.6 Cl-99 HCO3-24 AnGap-20 K-3.5 Cl-97 HCO3-24 AnGap-19 K-3.8 Cl-96 HCO3-21* AnGap-22* Glucose-NEG Ketone-80 Bilirub-NEG Urobiln-NEG pH-6.0 Leuks-NEG Epi-0 **FINAL REPORT ___. URINE CULTURE (Final ___: NO GROWTH. RADIOLOGY: * Phlegmon/ multiloculated fluid collection with surrounding. extensive inflammatory changes is indistinguishable from the distal portion of the appendix. Findings are concerning for perforated appendicitis. Possibility of underlying mass is difficult to exclude, particularly in this patient with history of melanoma. * Duodenal wall thickening may be inflammatory secondary to the. adjacent phlegmon. * Duodenum does not cross the midline, consistent with. intestinal malrotation. * Cholelithiasis. * Nonspecific bulbous appearance of the uncinate process of the. pancreas without discrete lesion identified. No pancreatic ductal dilatation. * Unorganized fluid/phlegmonous collection within the right. lower quadrant, surrounding the appendix, appears minimally enlarged since the reference study from ___. The findings favor ruptured appendicitis or a ruptured appendiceal mucocele. A neoplastic source relating to known history of melanoma would be atypical. Continued short-term imaging surveillance is recommended. * Congenital bowel malrotation, without volvulus or. obstruction. * Cholelithiasis. The remainder of the hospital course is summarized below: Neuro: The patient was alert and oriented throughout hospitalization; pain was initially managed with a IV dilaudid. He had left ankle pain and swelling consistent with gout that was managed with PO indomethacin.. CV: The patient remained stable from a cardiovascular standpoint; vital signs were routinely monitored. Pulmonary: The patient remained stable from a pulmonary standpoint. Good pulmonary toilet, early ambulation and incentive spirometry were encouraged throughout hospitalization. GI/GU/FEN: The patient was initially kept NPO. On HD3 he was given a clear liquid diet. On HD4 he was advanced to regular diet with good tolerability. Patient's intake and output were closely monitored. ID: The patient's fever curves were closely watched for signs of infection, of which there were none. He was initially given IV zosyn and transitioned to oral flagyl and ciprofloxacin upon discharge to complete a 2 week course of antibiotics. HEME: The patient's blood counts were closely watched for signs of bleeding, of which there were none. Prophylaxis: The patient received subcutaneous heparin and ___ dyne boots were used during this stay and was encouraged to get up and ambulate as early as possible. At the time of discharge, the patient was doing well, afebrile and hemodynamically stable. The patient was tolerating a diet, ambulating, voiding without assistance, and pain was well controlled. The patient received discharge teaching and follow-up instructions with understanding verbalized and agreement with the discharge plan. He was instructed to follow up with a colonoscopy outpatient in ___.<sep>A summary of the patient's past medical history is as follows: 1.right acoustic neuroma (deafness right ear) 2.s/p repair right biceps tendon rupture (___) 3.s/p right supraclavicular lymph node biopsy (___). PAST MEDICAL HISTORY: Stage IIIb melanoma diagnosed in ___ with findings of a positive right supraclavicular node, status post right anterior neck dissection revealing ___ additional positive nodes. He had adjuvant interferon therapy with Dr. ___ completed in ___, 36 weeks of this treatment. Bicep tendon repair. Acoustic neuroma followed with serial MRIs.<sep>Information on any allergies is detailed as follows: Codeine / Levaquin.<sep>Details on any major surgeries or invasive procedures are as follows: None.<sep>The service details are provided as follows: SURGERY.

Table 6: Input text from the EHR shown in Fig. 2 to generate the "Brief Hospital Course" section. The prompts used in both targets are highlighted in green and the prompt used only for "Brief Hospital Course" is highlighted in blue.

The patient's name is provided as follows: ___.<sep>Gender details are as follows: Male.<sep>The primary reason for the visit is summarized as follows: Abdominal pain.<sep>An overview of the current illness's history is provided as follows: Mr. ___ is a ___ male with PMHx significant for stage IIIb supraclavicular melanoma s/p supraclavicular and right anterior neck dissection and prostate cancer s/p radical prostatectomy, presenting with abdominal pain. The pain began two weeks ago and has been worsening in the last few days. His pain is localized to the right periumbilical region. He endorses having chills, inability to sleep and eat due to pain and a 6 lb weight loss in the past few days. He has been passing flatus and having loose and frequent non-bloody stools. He has also been having night sweats. He denies fever, nausea or vomiting. He was seen at ___ and transferred to ___ ED for further management after his CT abdomen showed a 5 x 6 x 7 cm right mid abdominal inflammatory phlegmon. His last colonoscopy was ___ years ago which revealed some polyps. He also complains of left ankle pain after a fall and has been taking ibuprofen.<sep>A summary of the patient's past medical history is as follows: 1.right acoustic neuroma (deafness right ear) 2.s/p repair right biceps tendon rupture (___) 3.s/p right supraclavicular lymph node biopsy (___). PAST MEDICAL HISTORY: Stage IIIb melanoma diagnosed in ___ with findings of a positive right supraclavicular node, status post right anterior neck dissection revealing ___ additional positive nodes. He had adjuvant interferon therapy with Dr. ___ completed in ___, 36 weeks of this treatment. Bicep tendon repair. Acoustic neuroma followed with serial MRIs.<sep>Information on any allergies is detailed as follows: Codeine / Levaquin.<sep>Details on any major surgeries or invasive procedures are as follows: None.<sep>Medications upon admission are detailed as follows: tadalafil (CIALIS) 5 mg daily PRN indomethacin 25 mg capsule TID.<sep>The service details are provided as follows: SURGERY.<sep>The final diagnosis at discharge is as follows: Perforated appendicitis.<sep>The disposition at discharge is provided as follows: Home.<sep>The patient's condition upon discharge is described as follows: Mental Status is Clear and coherent. Level of Consciousness is Alert and interactive. Activity Status is Ambulatory - Independent.<sep>Medications prescribed at discharge are as follows: * Acetaminophen 1000 mg PO TID. Do not exceed 4 grams/ 24 hours. * Ciprofloxacin HCl 500 mg PO Q12H. monitor for s/sx of allergic reaction RX *ciprofloxacin HCl 500 mg 1 tablet(s) by mouth twice a day. Disp #*20 Tablet Refills:*0 * Indomethacin 25 mg PO TID. RX *indomethacin 25 mg 1 capsule(s) by mouth three times a day. Disp #*42 Capsule Refills:*0 * MetroNIDAZOLE 500 mg PO Q8H. RX *metronidazole 500 mg 1 tablet(s) by mouth three times a day. Disp #*30 Tablet Refills:*0.

Table 7: Input text from the EHR in Fig. 2 to generate the "Discharge Instructions" section. The prompts used in both targets are highlighted in green and the prompts used only for "Discharge Instructions" are highlighted in orange.



(a) Brief Hospital Course

(b) Discharge Instructions

Figure 3: Histograms of the text length (in tokens) of the EHR and the input texts for the training and validation sets. The dashed line is the mean. The maximum text length is 1596 tokens, and see Table 12 in Appendix D for more details.

	Brief Hospital Course		Discharge instructions	
	EHR	Input	EHR	Input
Min	101	80	101	115
Max	8725	5664	8725	5774
Mean	1330	554	1330	639

Table 8: Statistical information (in tokens) for histograms in Fig. 3.

then provides examples and statistical information before and after preprocessing.

B.1 Extraction of simple sections

This section explains the process for extracting the "Sex", "Service", "Allergies", "Chief Complaint", and "Major Surgical or Invasive Procedure" sections.

To extract these sections, we used specific regular expressions such as `Sex: (\w+)\n`.

B.2 Extraction of complex sections

This section explains the process for extracting the "History of Present Illness", "Past Medical History", "Pertinent Results", "Medications on Admission", "Discharge Medications", "Discharge Disposition", "Discharge Diagnosis", and "Discharge Condition" sections.

We performed more detailed processing and pattern matching to efficiently extract the text of these sections. For example, for the "Discharge Condition" section, we used the regular expression `Discharge Diagnosis:\s*\n(.*?)(?=Discharge Condition:)` and it matches the diagnosis text up to the "Discharge Condition" section.

B.3 Detailed processing of each section

"Name". The patient's name is given as "___" and we used it directly.

"Sex". We converted "M" to "Male" and "F" to "Female".

"Pertinent Results". Timestamps in lines like "__ 08:00AM BLOOD __" were removed using regular expressions. In addition, list sections are converted to "*" format to maintain text consistency and clarity.

"Medications on Admission". List sections are converted to "*" format to maintain text consistency and clarity.

"Discharge Condition". We changed a colon in the extracted text to "is". For example, "Condition: Stable" is changed to "Condition is Stable".

"Discharge Medications". List sections are converted to "*" format to maintain text consistency and clarity.

B.4 Other processing

We ensure textual continuity by replacing line breaks with spaces and trimming excess spaces. In cases where no matching text is found, the default response is designated as "Unknown".

B.5 Examples of input text

Tables 6 and 7 show examples of input text. These examples illustrate that the ClinicalT5-large model is fine-tuned with different input text for each target discharge summary.

B.6 Statistical information

Fig. 3 shows histograms of the text length (in tokens) of the EHR and the input texts for the training and validation sets. Table 8 shows the statistical information for these histograms. As shown in Fig. 3 and Table 8, the preprocessing significantly reduces the length of the text.

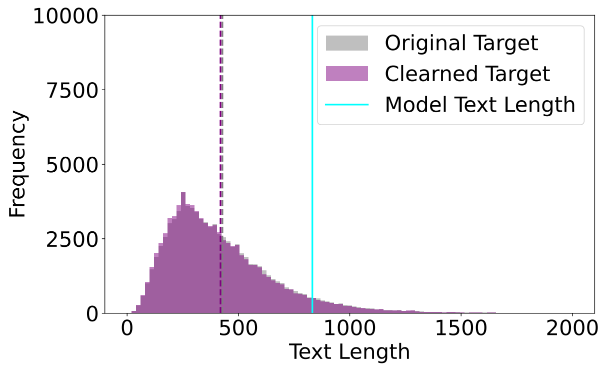
C Details of target text

C.1 Extraction and concatenation of segments

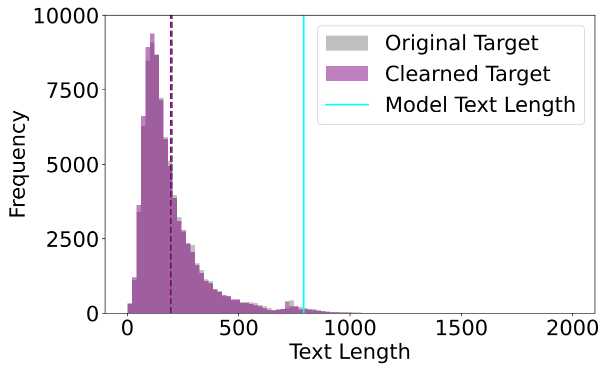
In the first process of segment extraction, we divide the text into segments based on blank lines and identify the distinct segments. We then remove spaces and line breaks from each segment and discard empty segments to retain only meaningful segments. Multiple consecutive spaces within each segment are replaced by a single space to improve readability. Finally, we reassemble the cleaned segments with line breaks to make them more suitable for training language models.

Text	Cleaned Text
Mr. ___ is a ___ yo M with medical history significant for \\ stage IIIb supraclavicular melanoma and prostate cancer admitted \\ to the Acute Care Surgery Service on ___ with worsening \\ abdominal pain, frequent stools, and subjective fevers. He was \\ transferred from ___ for further management with a CT \\ abdomen showing a 5 x 6 x 7 cm right mid abdominal inflammatory \\ phlegmon. He was admitted to the surgical floor for IV \\ antibitoics and further evaluation.\\ \\ Gastroenterology was consulted for duodenal thickening. Given \\ his current infection the wall thickening is likely secondary to \\ the infection. Repeat imaging was recommended to evaluate \\ evolution of the phlegmon as well as outpatient colonoscopy once \\ antibiotic treatment is complete. \\ \\ The remainder of the hospital course is summarized below:\\ Neuro: The patient was alert and oriented throughout \\ hospitalization; pain was initially managed with a IV dilaudid. \\ He had left ankle pain and swelling consistent with gout that \\ was managed with PO indomethacin.. \\ CV: The patient remained stable from a cardiovascular \\ standpoint; vital signs were routinely monitored.\\ Pulmonary: The patient remained stable from a pulmonary \\ standpoint. Good pulmonary toilet, early ambulation and \\ incentive spirometry were encouraged throughout hospitalization. \\ \\ GI/GU/FEN: The patient was initially kept NPO. On HD3 he was \\ given a clear liquid diet. On HD4 he was advanced to regular \\ diet with good tolerability. Patient's intake and output were \\ closely monitored\\ ID: The patient's fever curves were closely watched for signs of \\ infection, of which there were none. He was initially given IV \\ zosyn and transitioned to oral flagyl and ciprofloxacin upon \\ discharge to complete a 2 week course of antibiotics. \\ HEME: The patient's blood counts were closely watched for signs \\ of bleeding, of which there were none.\\ Prophylaxis: The patient received subcutaneous heparin and ___ \\ dyne boots were used during this stay and was encouraged to get \\ up and ambulate as early as possible.\\ \\ At the time of discharge, the patient was doing well, afebrile \\ and hemodynamically stable. The patient was tolerating a diet, \\ ambulating, voiding without assistance, and pain was well \\ controlled. The patient received discharge teaching and \\ follow-up instructions with understanding verbalized and \\ agreement with the discharge plan. He was instructed to follow \\ up with a colonoscopy outpatient in ___.	Mr. ___ is a ___ yo M with medical history significant for stage IIIb supraclavicular melanoma and prostate cancer admitted to the Acute Care Surgery Service on ___ with worsening abdominal pain, frequent stools, and subjective fevers. He was transferred from ___ for further management with a CT abdomen showing a 5 x 6 x 7 cm right mid abdominal inflammatory phlegmon. He was admitted to the surgical floor for IV antibitoics and further evaluation.\\ \\ Gastroenterology was consulted for duodenal thickening. Given his current infection the wall thickening is likely secondary to the infection. Repeat imaging was recommended to evaluate evolution of the phlegmon as well as outpatient colonoscopy once antibiotic treatment is complete.\\ \\ The remainder of the hospital course is summarized below:\\ Neuro: The patient was alert and oriented throughout hospitalization; pain was initially managed with a IV dilaudid. He had left ankle pain and swelling consistent with gout that was managed with PO indomethacin.. CV: The patient remained stable from a cardiovascular standpoint; vital signs were routinely monitored.\\ Pulmonary: The patient remained stable from a pulmonary standpoint. Good pulmonary toilet, early ambulation and incentive spirometry were encouraged throughout hospitalization.\\ \\ GI/GU/FEN: The patient was initially kept NPO. On HD3 he was given a clear liquid diet. On HD4 he was advanced to regular diet with good tolerability. Patient's intake and output were closely monitored\\ ID: The patient's fever curves were closely watched for signs of infection, of which there were none. He was initially given IV zosyn and transitioned to oral flagyl and ciprofloxacin upon discharge to complete a 2 week course of antibiotics. HEME: The patient's blood counts were closely watched for signs of bleeding, of which there were none.\\ Prophylaxis: The patient received subcutaneous heparin and ___ dyne boots were used during this stay and was encouraged to get up and ambulate as early as possible.\\ \\ At the time of discharge, the patient was doing well, afebrile and hemodynamically stable. The patient was tolerating a diet, ambulating, voiding without assistance, and pain was well controlled. The patient received discharge teaching and follow-up instructions with understanding verbalized and agreement with the discharge plan. He was instructed to follow up with a colonoscopy outpatient in ___.

Table 9: The text of the "Brief Hospital Course" section in Table 1 and its cleaned text by preprocessing. "\\" means line breaks.



(a) Brief Hospital Course



(b) Discharge Instructions

Figure 4: Histograms of the text length (in tokens) of the target texts before and after preprocessing for the training and validation sets. The dashed line is the mean. The maximum text length is 832 tokens for "Brief Hospital Course" and 792 tokens for "Discharge Instructions", see Table 12 in Appendix D for more details.

C.2 Examples of preprocessed target text

Tables 9 and 10 show examples of the target text before and after preprocessing. These examples illustrate that redundant line breaks are removed after preprocessing.

C.3 Statistical information

Fig. 4 shows histograms of the text length (in tokens) of the target texts before and after preprocessing for the training and validation sets. Table 11 shows the statistical information for these

Text	Cleaned Text
Dr. ___,\\ \\ You were admitted to the Acute Care Surgery Service on ___ \\ with abdominal pain. You had a CT scan of your abdomen that \\ showed likely a perforated appendicitis. You were given IV \\ antibiotics and had improvement in your symptoms. An attempt was \\ made to drain the infection but it is not amenable to a drain at \\ this time. You were transitioned to oral antibiotics with \\ continued good effect.\\ \\ While in the hospital you had a flair up of gout in your left \\ ankle. You were given indomethacin with improvement in your \\ symptoms.\\ \\ You are now doing better, tolerating a regular diet, and ready \\ to be discharged to home to continue your recovery.\\ \\ Please note the following discharge instructions:\\ \\ Please call your doctor or nurse practitioner or return to the \\ Emergency Department for any of the following:\\ *You experience new chest pain, pressure, squeezing or \\ tightness.\\ *New or worsening cough, shortness of breath, or wheeze.\\ *If you are vomiting and cannot keep down fluids or your \\ medications.\\ *You are getting dehydrated due to continued vomiting, diarrhea, \\ or other reasons. Signs of dehydration include dry mouth, rapid \\ heartbeat, or feeling dizzy or faint when standing.\\ *You see blood or dark/black material when you vomit or have a \\ bowel movement.\\ *You experience burning when you urinate, have blood in your \\ urine, or experience a discharge.\\ *Your pain in not improving within ___ hours or is not gone \\ within 24 hours. Call or return immediately if your pain is \\ getting worse or changes location or moving to your chest or \\ back.\\ *You have shaking chills, or fever greater than 101.5 degrees \\ Fahrenheit or 38 degrees Celsius.\\ *Any change in your symptoms, or any new symptoms that concern \\ you.\\ \\ Please resume all regular home medications, unless specifically \\ advised not to take a particular medication. Also, please take \\ any new medications as prescribed.\\ \\ Please get plenty of rest, continue to ambulate several times \\ per day, and drink adequate amounts of fluids.	Dr. ___,\\ \\ You were admitted to the Acute Care Surgery Service on ___ with abdominal pain. You had a CT scan of your abdomen that showed likely a perforated appendicitis. You were given IV antibiotics and had improvement in your symptoms. An attempt was made to drain the infection but it is not amenable to a drain at this time. You were transitioned to oral antibiotics with continued good effect.\\ \\ While in the hospital you had a flair up of gout in your left ankle. You were given indomethacin with improvement in your symptoms.\\ \\ You are now doing better, tolerating a regular diet, and ready to be discharged to home to continue your recovery.\\ \\ Please note the following discharge instructions:\\ \\ Please call your doctor or nurse practitioner or return to the Emergency Department for any of the following:\\ *You experience new chest pain, pressure, squeezing or tightness.\\ *New or worsening cough, shortness of breath, or wheeze.\\ *If you are vomiting and cannot keep down fluids or your medications.\\ *You are getting dehydrated due to continued vomiting, diarrhea, or other reasons. Signs of dehydration include dry mouth, rapid heartbeat, or feeling dizzy or faint when standing.\\ *You see blood or dark/black material when you vomit or have a bowel movement.\\ *You experience burning when you urinate, have blood in your urine, or experience a discharge.\\ *Your pain in not improving within ___ hours or is not gone within 24 hours. Call or return immediately if your pain is getting worse or changes location or moving to your chest or back.\\ *You have shaking chills, or fever greater than 101.5 degrees Fahrenheit or 38 degrees Celsius.\\ *Any change in your symptoms, or any new symptoms that concern you.\\ \\ Please resume all regular home medications, unless specifically advised not to take a particular medication. Also, please take any new medications as prescribed.\\ \\ Please get plenty of rest, continue to ambulate several times per day, and drink adequate amounts of fluids.

Table 10: The text of the "Discharge Instructions" section in Table 1 and its cleaned text by preprocessing. "\\" means line breaks.

	Brief Hospital Course		Discharge instructions	
	Original Target	Cleaned Target	Original Target	Cleaned Target
Min	2	1	10	10
Max	4614	4452	5025	4861
Mean	428	419	201	195

Table 11: Statistical information (in tokens) for histograms in Fig. 4.

	Brief Hospital Course	Discharge instructions
Input Text	1596	
Generated Text	832	792
Text	832	792

Table 12: Maximum text length (tokens).

histograms. As shown in Fig. 4 and Table 11, the preprocessing slightly reduces the length of the text.

D Details of fine-tuning

We used Pytorch (Paszke et al., 2019) and hugging-face transformers (Wolf et al., 2020) to implement and fine-tune our models. We also use peft (Man-

Batch size	2
Epochs	4
Learning rate	1e-4
Precision setting	FP16
Weight decay	0.01

Table 13: Hyperparameters for fine-tuning.

Dropout probability	0.05
Rank	4
Target modules	Query & Value
α	16

Table 14: Hyperparameters for LoRA.

Min length	10
Num beams	4
Do sample	True
Length penalty	1.1
No repeat n-gram size	4

Table 15: Hyperparameters to generate each target discharge summary.

grulkar et al., 2022) for LoRA.

Table 12 shows the text length (in tokens) used by our models. Table 13 shows the hyperparameters used for fine tuning. Table 14 shows the hyperparameters used for LoRA. Table 15 shows the hyperparameters to generate each target discharge summary.