

MedDecXtract: A Clinician-Support System for Extracting, Visualizing, and Annotating Medical Decisions in Clinical Notes

Mohamed Elgaar

MELGAAR@CS.UML.EDU and Hadi Amiri

HADI@CS.UML.EDU

University of Massachusetts Lowell, USA

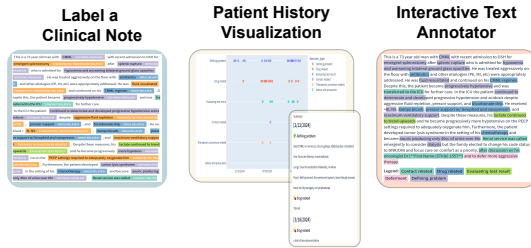


Figure 1: Overview of MedDecXtract functionalities: 1) **Labeling Clinical Notes**: Highlights key medical decisions using color-coded labels for different decision categories. 2) **Patient History Visualization**: Aggregates multiple clinical notes into a timeline to visualize decision sequences over time. 3) **Interactive Text Annotation**: Allows manual labeling of medical decisions with support for pseudo-annotations to expedite the process.

1. Introduction

Clinical narratives, such as discharge summaries, are rich sources of information regarding patient care and clinical decision-making. These texts contain crucial details about diagnoses, treatments, medications, and follow-up plans. However, manually extracting this information is time-consuming and labor-intensive. Automating this process is essential for enhancing clinical decision support.

Previous work has focused on extracting medical concepts, drugs, diseases, and other entities (Nye et al., 2018; Lehman et al., 2019; Patel et al., 2018). While valuable, these efforts do not directly address the extraction of med-

Table 1: DICTUM Categories

Category	Description
Contact related	Admit, discharge
Gathering info	Ordering test, consulting
Defining problem	Diagnosis, prognosis
Treatment goal	Quant./Qual. Goal
Drug related	Start, stop, alter
Therapeutic procedure	Start, stop, alter
Evaluating test	Positive, negative
Deferment	Transfer, wait
Advice/precaution	Advice or precaution
Legal/insurance	Sick leave, refund

ical decisions. **MedDecXtract** addresses this gap by employing a transformer-based model for span extraction and classification, specifically using RoBERTa, to identify medical decisions as defined by the DICTUM taxonomy (Ofstad et al., 2016). This approach provides a more granular and actionable level of information extraction from clinical notes, intended for use by clinicians and researchers. The MedDec dataset (Elgaar et al., 2024) offers a comprehensive resource of annotated medical decisions within discharge summaries, targeting a broad patient population. MedDecXtract provides an interactive interface for exploring and annotating these decisions, streamlining the creation of high-quality labeled data.¹

1. Demo available at <https://mohdelgaar-clinical-decisions.hf.space>.

2. Method

MedDecXtract employs a fine-tuned RoBERTa (Liu et al., 2019) model for sequence labeling, using token classification to extract and classify decision spans according to DICTUM taxonomy categories, shown in Table 1. The MedDec dataset (Elgaar et al., 2024), sourced from the MIMIC-III clinical database (Pollard and Johnson III, 2016), provides annotated discharge summaries with ten types of medical decisions. The tool offers three primary functionalities:

1. **Labeling Clinical Notes:** Users input a clinical note to receive highlighted medical decisions, categorized into predefined types.
2. **Patient History Visualization:** Users can upload multiple notes for a single patient to visualize the timeline of medical decisions, aiding in understanding the decision-making process over time.
3. **Interactive Text Annotation:** An intuitive interface for manually annotating medical decisions, reducing manual effort and improving dataset quality.

The deployment pipeline uses Gradio (Abid et al., 2019) to create an interactive web interface, hosted on Hugging Face Spaces (Face, 2024), enabling real-time interaction and visualization for clinicians and researchers.

3. Results

MedDecXtract is hosted on Hugging Face Spaces, demonstrating its potential to support clinicians and researchers by automating the annotation of medical decisions. During development, the tool achieved high accuracy in extracting and classifying medical decisions on the MedDec dataset. Although specific deployment metrics are not yet available, the tool is designed to facilitate the annotation of newer clinical note corpora, such as MIMIC-IV, aiding in research tasks like studying calibration shift/drift. By integrating into annotation workflows, MedDecXtract can significantly reduce manual effort and improve the quality of labeled datasets, assisting clinicians in quickly understanding patient histories and discharge summaries.

4. Discussion

Developing MedDecXtract presented challenges, such as ensuring the accuracy of extracted decisions and integrating the tool into existing clinical workflows. Future improvements will focus on enhancing the model’s accuracy and expanding its applicability to other clinical settings. Additionally, further deployment and user testing will be crucial in understanding its impact and refining its features to better meet user needs. The potential to annotate more data and study calibration shift/drift in newer datasets like MIMIC-IV presents exciting opportunities for future research and development.

References

- Abubakar Abid, Ali Abdalla, Ali Abid, Dawood Khan, Abdulrahman Alfozan, and James Zou. Gradio: Hassle-free sharing and testing of ml models in the wild. *arXiv preprint arXiv:1906.02569*, 2019.
- Mohamed Elgaar, Jiali Cheng, Nidhi Vakil, Hadi Amiri, and Leo Anthony Celi. MedDec: A dataset for extracting medical decisions from discharge summaries. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Findings of the Association for Computational Linguistics ACL 2024*, pages 16442–16455, Bangkok, Thailand and virtual meeting, August 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.findings-acl.975. URL <https://aclanthology.org/2024.findings-acl.975>.
- Hugging Face. Hugging face, 2024. URL <https://huggingface.co>.
- Eric Lehman, Jay DeYoung, Regina Barzilay, and Byron C. Wallace. Inferring which medical treatments work from reports of clinical trials. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *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 3705–3717, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1371. URL <https://aclanthology.org/N19-1371>.

- 125 Yinhan Liu, Myle Ott, Naman Goyal, Jingfei
 126 Du, Mandar Joshi, Danqi Chen, Omer Levy,
 127 Mike Lewis, Luke Zettlemoyer, and Veselin
 128 Stoyanov. Roberta: A robustly optimized
 129 bert pretraining approach. *ArXiv preprint*,
 130 abs/1907.11692, 2019. URL <https://arxiv.org/abs/1907.11692>.
 131 [org/abs/1907.11692](https://arxiv.org/abs/1907.11692).
- 132 Benjamin Nye, Junyi Jessy Li, Roma Patel, Yin-
 133 fei Yang, Iain Marshall, Ani Nenkova, and By-
 134 ron Wallace. A corpus with multi-level an-
 135 notations of patients, interventions and out-
 136 comes to support language processing for med-
 137 ical literature. In Iryna Gurevych and Yusuke
 138 Miyao, editors, *Proceedings of the 56th An-
 139 nual Meeting of the Association for Computa-
 140 tional Linguistics (Volume 1: Long Papers)*,
 141 pages 197–207, Melbourne, Australia, July
 142 2018. Association for Computational Linguis-
 143 tics. doi: 10.18653/v1/P18-1019. URL [https:
 144 //aclanthology.org/P18-1019](https://aclanthology.org/P18-1019).
- 145 Eirik H Ofstad, Jan C Frich, Edvin Schei,
 146 Richard M Frankel, and Pål Gulbrandsen.
 147 What is a medical decision? a taxonomy based
 148 on physician statements in hospital encounters:
 149 a qualitative study. *BMJ open*, 6(2):e010098,
 150 2016. URL [https://pubmed.ncbi.nlm.nih.
 151 gov/26868946/](https://pubmed.ncbi.nlm.nih.gov/26868946/).
- 152 Pinal Patel, Disha Davey, Vishal Panchal, and
 153 Parth Pathak. Annotation of a large clini-
 154 cal entity corpus. In Ellen Riloff, David Chi-
 155 ang, Julia Hockenmaier, and Jun’ichi Tsu-
 156 jii, editors, *Proceedings of the 2018 Confer-
 157 ence on Empirical Methods in Natural Lan-
 158 guage Processing*, pages 2033–2042, Brussels,
 159 Belgium, October-November 2018. Association
 160 for Computational Linguistics. doi: 10.18653/
 161 v1/D18-1228. URL [https://aclanthology.
 162 org/D18-1228](https://aclanthology.org/D18-1228).
- 163 Tom J Pollard and AEW Johnson III. The
 164 mimic iii clinical database, version 1.4.
 165 *The MIMIC-III Clinical Database. Phys-
 166 ioNet*, 2016. URL [https://physionet.org/
 167 content/mimiciii/1.4/](https://physionet.org/content/mimiciii/1.4/).