PhysioNet

Search

MIMIC-III database

Sort by   Relevance

Resource type   4 selected
- ☑ Data
- ☑ Software
- ☑ Challenge
- ☑ Model

Apply

Resources

🗄 Database   🔒 Credentialed Access

[EchoNotes Structured Database derived from MIMIC-III (ECHO-NOTE2NUM)](#)

Gloria Hyunjung Kwak, Dana Moukheiber, Mira Moukheiber, Lama Moukheiber, Sulaiman Moukheiber, Neel Butala, Leo Anthony Celi, Christina Chen

A structured echocardiogram database derived from 43,472 observational notes obtained during echocardiogram studies conducted in the intensive care unit at the Beth Israel Deaconess Medical Center between 2001 and 2012.

Published: Feb. 23, 2024. Version: 1.0.0

🗄 Database   🔒 Credentialed Access

Annotation dataset of problematic opioid use and related contexts from MIMIC-III Critical Care Database discharge summaries

Melissa Poulsen, Vanessa Troiani, Philip Freda, Danielle Mowery, Anahita Davoudi

The database contains a corpus of annotated data from the MIMIC-III Critical Care Database

from a study that aimed to develop and apply an annotation schema to characterize opioid use disorder and related contextual factors.

opioid use disorder   substance use

natural language processing

clinical notes

Published: Feb. 8, 2023. Version: 1.0.0

🗄 Database   🔒 Credentialed Access

[Annotation dataset of social determinants of health from MIMIC-III Clinical Care Database](#)

Marco Guevara, Shan Chen, Spencer Thomas, Danielle Bitterman

Annotation dataset of social determinants of health from MIMC-III Clinical Care Database notes.

natural language processing

social determinants of health

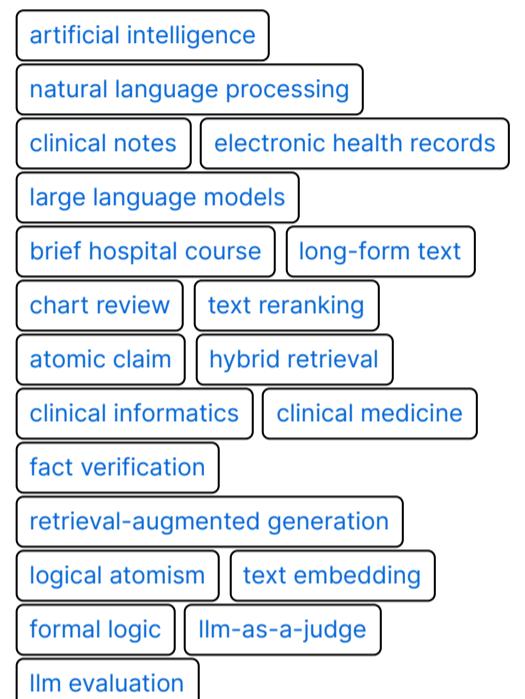Published: Jan. 24, 2024. Version: 1.0.1

🗄 Database   🔒 Credentialed Access

[MIMIC-III-Ext-VeriFact-BHC: Labeled Propositions From Brief Hospital Course Summaries for Long-form Clinical Text Evaluation](#)

Philip Chung, Akshay Swaminathan, Alex Goodell, Yeasul Kim, Momsen Reincke, Lichy Han, Ben Deverett, Mohammad Amin Sadeghi, Abdel badih El Ariss, Marc Ghanem, David Seong, Andrew Lee, Caitlin Coombes, Brad Bradshaw, Mahir

Sufian, Hyo Jung Hong, Teresa Nguyen, Mohammad Rasouli, Komal Kamra, Mark Burbridge, James McAvoy, Roya Saffary, Stephen Parnell Ma, Dev Dash, James Xie, Ellen Wang, Cliff Schmiesing, Nigam Shah, Nima Aghaeepour

A clinician-labeled dataset for fact-checking long-form clinical text against patient EHRs. The dataset contains LLM-written and human-written Brief Hospital Course summaries decomposed to atomic claim and sentence propositions with annotations.
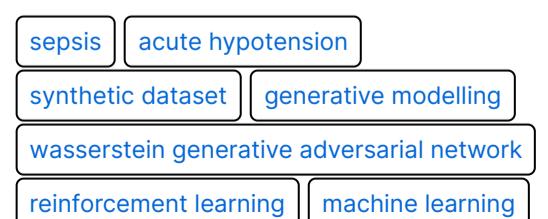
artificial intelligence

natural language processing

clinical notes   electronic health records

large language models

brief hospital course   long-form text

chart review   text reranking

atomic claim   hybrid retrieval

clinical informatics   clinical medicine

fact verification

retrieval-augmented generation

logical atomism   text embedding

formal logic   llm-as-a-judge

llm evaluation

Published: April 9, 2025. Version: 1.0.0

Database   🔒 Credentialed Access

[Synthetic Acute Hypotension and Sepsis Datasets Based on MIMIC-III and Published as Part of the Health Gym Project](#)

Nicholas Kuo, Simon Finfer, Louisa Jorm, Sebastiano Barbieri

This repository hosts the original Health Gym datasets of Acute Hypotension and Sepsis

sepsis   acute hypotension

synthetic dataset   generative modelling

wasserstein generative adversarial network

reinforcement learning   machine learning

Published: Feb. 23, 2022. Version: 1.0.0

Model | Credentialed Access

[Transformer models trained on MIMIC-III to generate synthetic patient notes](#)

Ali Amin-Nejad, Julia Ive, Sumithra Velupillai

Machine learning models that have been trained using MIMIC-III to enable the creation of synthetic discharge summaries.

Published: May 27, 2020. Version: 1.0.0

Database | Credentialed Access

[Nosocomial Risk Datasets from MIMIC-III](#)

Travis Goodwin

Text-based Longitudinal Data for Predicting Nosocomial Disease Risk as used by CANTRIP.

pressure injury | risk prediction
acute kidney injury | anemia
forecasting | natural language processing
deep learning

Published: Sept. 15, 2022. Version: 1.0

Database | Credentialed Access

[MIMIC-III and eICU-CRD: Feature Representation by FIDDLE Preprocessing](#)

Shengpu Tang, Parmida Davarmanesh, Yanmeng Song, Danai Koutra, Michael Sjoding, Jenna Wiens

Features and labels from MIMIC-III and eICU-CRD produced by FIDDLE, an EHR preprocessing pipeline.

preprocessing | electronic health record | machine learning

Published: April 28, 2021. Version: 1.0.0

Database | 🔒 Credentialed Access

Northwestern ICU (NWICU) database

Dana Moukheiber, William Temps, Bhadrappa Molgi, Yikuan Li, Alice Lu, Prasanth Nannapaneni, Abdulrahman Chahin, Sicheng Hao, Felipe Torres Fabregas, Leo Anthony Celi, Adrian Wong, Maxwell Lloyd, Xavier Borrat Frigola, Hyung-Chul Lee, Daniel Schneider, Tom Pollard, Yuan Luo, Abel Kho, Roger Mark

A freely available COVID-rich ICU database comprising de-identified health-related data from Northwestern Memorial Health Center (NHMC).

Published: Nov. 19, 2024. Version: 0.1.0

Database | 🔒 Credentialed Access

BOLD, a blood-gas and oximetry linked dataset

João Matos, Tristan Struja, Jack Gallifant, Luis Filipe Nakayama, Marie Charpignon, Xiaoli Liu, Jaime dos Santos Cardoso, Leo Anthony Celi, An Kwok Wong

An open-source pulse oximetry and arterial blood gas dataset, derived from MIMIC-III, MIMIC-IV, and eICU-CRD

pulse oximetry | intensive care unit | health equity | electronic health records

Published: Nov. 8, 2023. Version: 1.0

MIT Laboratory for Computational Physiology

National Institute of Biomedical Imaging and Bioengineering (NIBIB) under NIH grant number R01EB030362

Navigation

Discover Data

Share Data

About

News

Explore

Data

Software

Tutorials

Challenges