MIMIC - III
Medical Information Mart for Intensive Care

Katarzyna Kobylińska

- https://mimic.physionet.org/

- Kody do ekstrakcji danych, MIT lab, czy MIMIC-extract

Hospital

ICU

MICU   SICU   CCU   CSRU   NICU

Bedside monitoring
- Vital signs
- Waveforms
- Trends
- Alarms

Chart
- Fluids
- Medications
- Progress notes

Tests
- Laboratory
- Microbiology

Orders
- Provider order entry (POE)

Billing
- ICD9
- DRG
- Procedures (CPT)

Demographics
- Admission/discharge dates
- Date of birth/death
- Religion/ethnicity/marital status

Notes and reports
- Discharge summaries
- Radiology (X-ray, CT, MRI, Ultrasound)
- Cardiology (ECHO, ECG)

External

Social Security Death Index

Data archive

De-identification

Date shifting

Format conversion

MIMIC-III Database

User feedback and corrections

Dostęp do bazy

- Kurs dotyczący ochrony wrażliwych danych dotyczących zdrowia: Health Insurance Portability and Accountability Act (HIPAA)
- Zgody na odpowiednie używanie danych, m.in zakaz dzielenia się danymi, odpowiednie cytowanie w publikacjach, zakaz identyfikacji pacjentów.
- Dostęp otrzymujemy w przeciągu co najmniej tygodnia
- Dostęp do instrukcji w jaki sposób ściągnąć bazę danych z PhysioNetWorks

Zalety MIMIC III

1. Dostępność, możliwość publikacji, reprodukcja kodów:

- Johnson, A., Pollard, T., Shen, L. *et al.* MIMIC-III, a freely accessible critical care database. *Sci Data* **3,** 160035 (2016).

  *https://doi.org/10.1038/sdata.2016.35database. Sci Data **3,** 160035 (2016).*

- Jedyna dostępna baza dotycząca pacjentów na OIOM
- Analizy nie są ograniczone (gdy mamy dostęp do bazy)

2. Możliwość modelowania różnych problemów na tych samych danych

3. Luka XAI + MIMIC III

4. Popularna w zastosowaniach ML w MED

Charakterystyka danych

Critical care unit	CCU	CSRU	MICU	SICU	TSICU	Total
Distinct patients, no. (% of total admissions)	5,674 (14.7%)	8,091 (20.9%)	13,649 (35.4%)	6,372 (16.5%)	4,811 (12.5%)	38,597 (100%)
Hospital admissions, no. (% of total admissions)	7,258 (14.6%)	9,156 (18.4%)	19,770 (39.7%)	8,110 (16.3%)	5,491 (11.0%)	49,785 (100%)
Distinct ICU stays, no. (% of total admissions)	7,726 (14.5%)	9,854 (18.4%)	21,087 (39.5%)	8,891 (16.6%)	5,865 (11.0%)	53,423 (100%)
Age, years, median (Q1-Q3)	70.1 (58.4–80.5)	67.6 (57.6–76.7)	64.9 (51.7–78.2)	63.6 (51.4–76.5)	59.9 (42.9–75.7)	65.8 (52.8–77.8)
Gender, male, % of unit stays	4,203 (57.9%)	6,000 (65.5%)	10,193 (51.6%)	4,251 (52.4%)	3,336 (60.7%)	27,983 (55.9%)
ICU length of stay, median days (Q1-Q3)	2.2 (1.2–4.1)	2.2 (1.2–4.0)	2.1 (1.2–4.1)	2.3 (1.3–4.9)	2.1 (1.2–4.6)	2.1 (1.2–4.6)
Hospital length of stay, median days (Q1-Q3)	5.8 (3.1–10.0)	7.4 (5.2–11.4)	6.4 (3.7–11.7)	7.9 (4.4–14.2)	7.4 (4.1–13.6)	6.9 (4.1–11.9)
ICU mortality, percent of unit stays	685 (8.9%)	353 (3.6%)	2,222 (10.5%)	813 (9.1%)	492 (8.4%)	4,565 (8.5%)
Hospital mortality, percent of unit stays	817 (11.3%)	424 (4.6%)	2,859 (14.5%)	1,020 (12.6%)	628 (11.4%)	5,748 (11.5%)

CCU is Coronary Care Unit; CSRU is Cardiac Surgery Recovery Unit; MICU is Medical Intensive Care Unit; SICU is Surgical Intensive Care Unit; TSICU is Trauma Surgical Intensive Care Unit.

Johnson, A., Pollard, T., Shen, L. *et al.* MIMIC-III, a freely accessible critical care database. *Sci Data* **3,** 160035 (2016). https://doi.org/10.1038/sdata.2016.35

Klasy zmiennych

Class of data	Description
Billing	Coded data recorded primarily for billing and administrative purposes. Includes Current Procedural Terminology (CPT) codes, Diagnosis-Related Group (DRG) codes, and International Classification of Diseases (ICD) codes.
Descriptive	Demographic detail, admission and discharge times, and dates of death.
Dictionary	Look-up tables for cross referencing concept identifiers (for example, International Classification of Diseases (ICD) codes) with associated labels.
Interventions	Procedures such as dialysis, imaging studies, and placement of lines.
Laboratory	Blood chemistry, hematology, urine analysis, and microbiology test results.
Medications	Administration records of intravenous medications and medication orders.
Notes	Free text notes such as provider progress notes and hospital discharge summaries.
Physiologic	Nurse-verified vital signs, approximately hourly (e.g., heart rate, blood pressure, respiratory rate).
Reports	Free text reports of electrocardiogram and imaging studies.

Johnson, A., Pollard, T., Shen, L. *et al.* MIMIC-III, a freely accessible critical care database. *Sci Data* **3,** 160035 (2016). https://doi.org/10.1038/sdata.2016.35

Przykładowe dane

1. Patient level data SUBJECT_ID
2. Hospital level data HADM_ID
3. ICU level data ICUSTAY_ID

```
  ROW_ID SUBJECT_ID HADM_ID ICUSTAY_ID          CHARTTIME ITEMID VALUE VALUEUOM           STORETIME  CGID STOPPED NEWBOTTLE ISERROR
1    344      21219  177991     225765 2142-09-08 10:00:00  40055   200        ml 2142-09-08 12:08:00 17269      NA        NA      NA
2    345      21219  177991     225765 2142-09-08 12:00:00  40055   200        ml 2142-09-08 12:08:00 17269      NA        NA      NA
3    346      21219  177991     225765 2142-09-08 13:00:00  40055   120        ml 2142-09-08 13:39:00 17269      NA        NA      NA
4    347      21219  177991     225765 2142-09-08 14:00:00  40055   100        ml 2142-09-08 16:17:00 17269      NA        NA      NA
5    348      21219  177991     225765 2142-09-08 16:00:00  40055   200        ml 2142-09-08 16:17:00 17269      NA        NA      NA
6    349      21219  177991     225765 2142-09-08 17:00:00  40055   100        ml 2142-09-08 17:16:00 17269      NA        NA      NA
7    350      21219  177991     225765 2142-09-08 18:00:00  40055   120        ml 2142-09-08 18:56:00 17269      NA        NA      NA
8    351      21219  177991     225765 2142-09-08 20:00:00  40055    60        ml 2142-09-08 19:39:00 19525      NA        NA      NA
```

```
ROW_ID SUBJECT_ID HADM_ID          ADMITTIME           DISCHTIME DEATHTIME ADMISSION_TYPE        ADMISSION_LOCATION       DISCHARGE_LOCATION INSURANCE LANGUAGE         RELIGION MARITAL_STATUS
    21         22  165315 2196-04-09 12:26:00 2196-04-10 15:54:00             EMERGENCY    EMERGENCY ROOM ADMIT DISC-TRAN CANCER/CHLDRN H   Private               UNOBTAINABLE        MARRIED
    22         23  152223 2153-09-03 07:15:00 2153-09-08 19:10:00              ELECTIVE PHYS REFERRAL/NORMAL DELI        HOME HEALTH CARE  Medicare                   CATHOLIC        MARRIED
    23         23  124321 2157-10-18 19:34:00 2157-10-25 14:00:00             EMERGENCY TRANSFER FROM HOSP/EXTRAM        HOME HEALTH CARE  Medicare     ENGL          CATHOLIC        MARRIED
    24         24  161859 2139-06-06 16:14:00 2139-06-09 12:48:00             EMERGENCY TRANSFER FROM HOSP/EXTRAM                    HOME   Private          PROTESTANT QUAKER         SINGLE
```

```
ROW_ID ICD9_CODE        SHORT_TITLE
   174     01166  TB pneumonia-oth test
   175     01170  TB pneumothorax-unspec
   176     01171  TB pneumothorax-no exam
   177     01172 TB pneumothorx-exam unkn
   178     01173 TB pneumothorax-micro dx
   179     01174  TB pneumothorax-cult dx
```

Ekstrakcja danych

- MIT lab: Massachusetts Institute of Technology (MIT)'s Laboratory for Computational Physiology

- https://github.com/MIT-LCP/mimic-code/tree/master/buildmimic

100 GB wolnej przestrzeni na dysku

Ciekawe artykuły dotyczące MIMIC:

MIMIC-III, a freely accessible critical care database

Alistair E.W. Johnson, Tom J. Pollard ✉, Lu Shen, Li-wei H. Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi & Roger G. Mark

Abstract

MIMIC-III ('Medical Information Mart for Intensive Care') is a large, single-center database comprising information relating to patients admitted to critical care units at a large tertiary care hospital. Data includes vital signs, medications, laboratory measurements, observations and notes charted by care providers, fluid balance, procedure codes, diagnostic codes, imaging reports, hospital length of stay, survival data, and more. The database supports applications including academic and industrial research, quality improvement initiatives, and higher education coursework.

Design Type(s)	data integration objective
Measurement	Demographics • clinical measurement • intervention • Billing • Medical History

Johnson, A., Pollard, T., Shen, L. *et al.* MIMIC-III, a freely accessible critical care database. *Sci Data* **3,** 160035 (2016). https://doi.org/10.1038/sdata.2016.35

Ciekawe artykuły dotyczące MIMIC:

Predictive modeling in urgent care: a comparative study of machine learning approaches
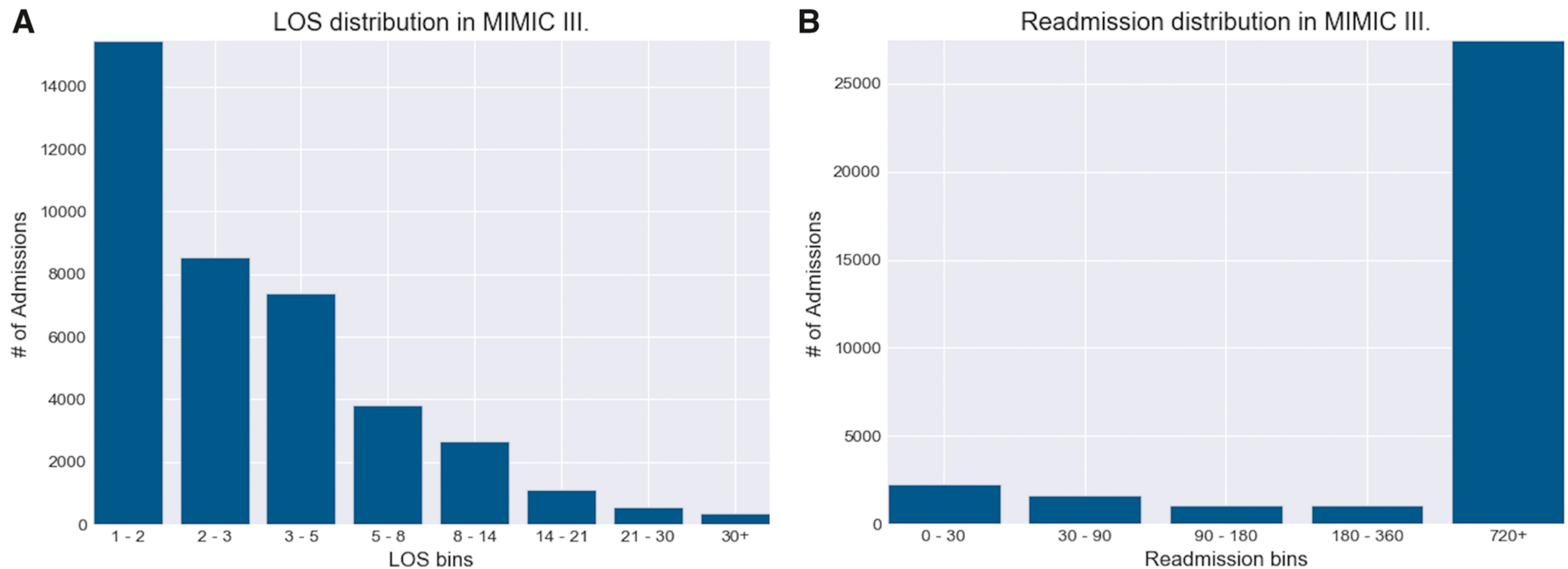

Figure 1. Distribution of length of stays (LOS) and readmission in MIMIC-III.

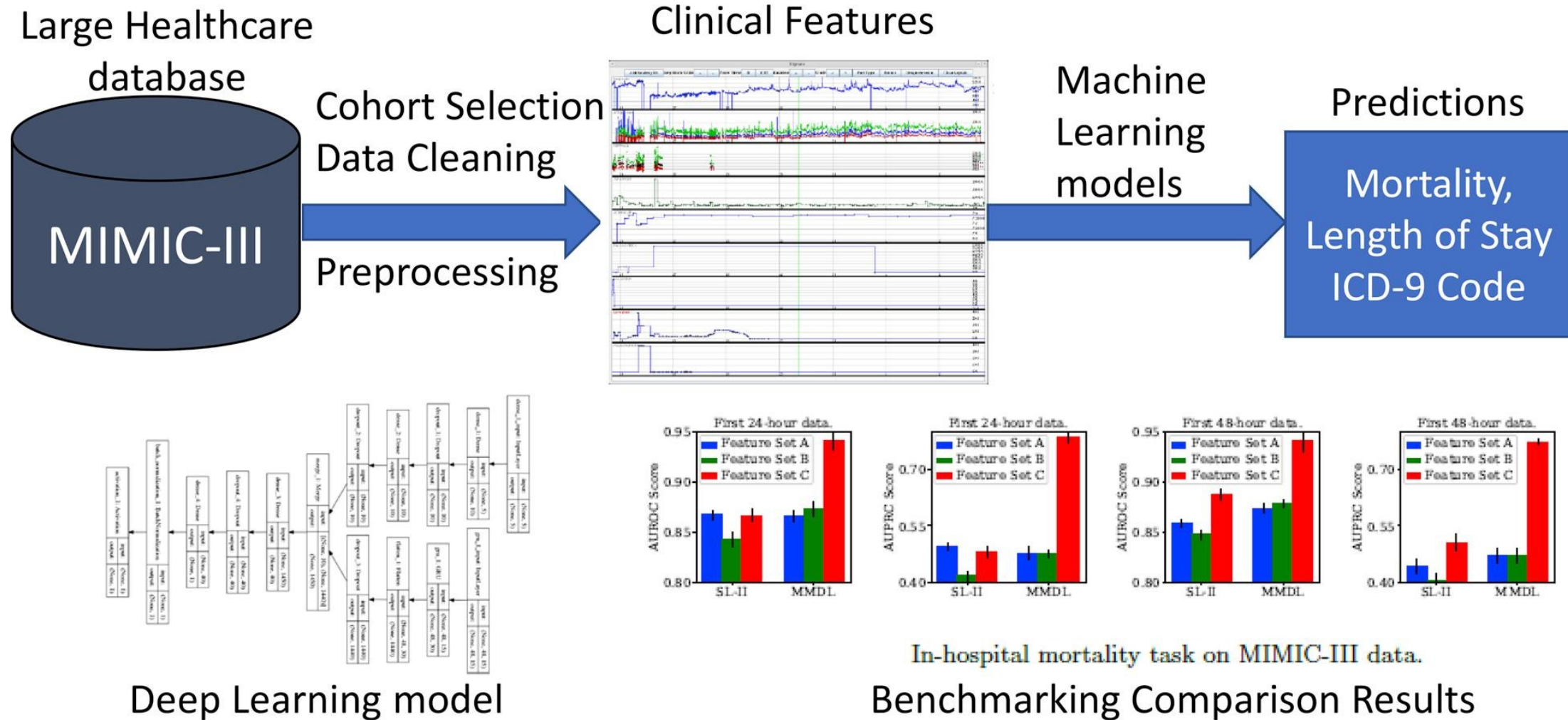# Ciekawe artykuły dotyczące MIMIC:

Predictive modeling in urgent care: a comparative study of machine learning approaches

Task: prawdopodobieństwo ponownego pobytu na oddziale

Rank	Model	AUC	F1	Sn	Sp	P-value
Classic models						
1	RF w/ w48	0.582 (0.0067)	0.122 (0.0025)	0.601 (0.02)	0.563 (0.0086)	.0387
2	LR w/ w2v	0.577 (0.0067)	0.123 (0.0023)	0.574 (0.031)	0.592 (0.0211)	.0469
3	RF w/ 48h	0.577 (0.009)	0.121 (0.003)	0.571 (0.021)	0.583 (0.004)	.0657
Sequential models						
1	LSTM w/ x19 + h2v	0.580 (0.00914)	0.112 (0.0043)	0.548 (0.0192)	0.565 (0.0206)	.0606
2	LSTM w/ x19	0.554 (0.00648)	0.108 (0.0028)	0.538 (0.0168)	0.554 (0.0214)	.107
3	LSTM w/ w2v	0.552 (0.0154)	0.107 (0.0038)	0.567 (0.0404)	0.524 (0.0272)	.199

Ciekawe artykuły dotyczące MIMIC:

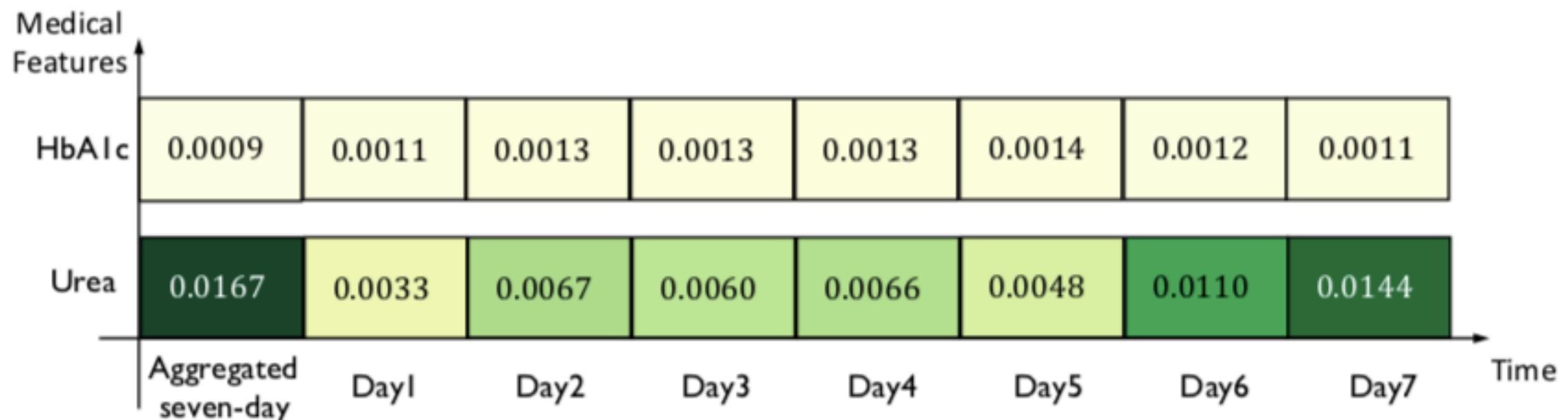Benchmarking deep learning models on large healthcare datasets

Ciekawe artykuły dotyczące XAI + MIMIC:

**TRACER**: A Framework for Facilitating Accurate and Interpretable Analytics for High Stakes Applications, https://arxiv.org/abs/2003.12012, ACM SIGMOD 2020

Designing Theory-Driven User-Centric Explainable AI
Conference on Human Factors in Computing Systems, May 4–9, 2019

Ciekawe artykuły dotyczące XAI + MIMIC:

TRACER: A Framework for Facilitating Accurate and Interpretable Analytics for High Stakes Applications

TRACER to provide accurate and **inTerpRetAble Clinical dEcision suppoRt to doctors and practitioners of other domains**

TRACER: A Framework for Facilitating Accurate and Interpretable Analytics for High Stakes Applications, https://arxiv.org/abs/2003.12012

Ciekawe artykuły dotyczące XAI + MIMIC:

TRACER: A Framework for Facilitating Accurate and Interpretable Analytics for High Stakes Applications

TRACER:

- **Istniejące metody nadmiernie zależne od przedziałów czasowych, bez różnicowania ważności cech niezmiennych i zmiennych w czasie**

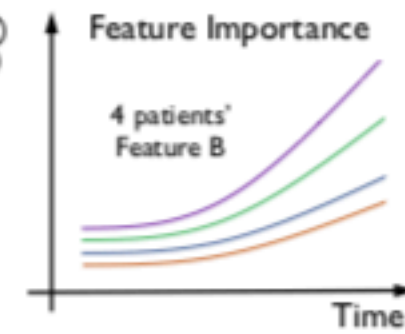- **Zapewnienie dokładnego i możliwego do interpretacji wsparcia w podejmowaniu decyzji lekarzom.**

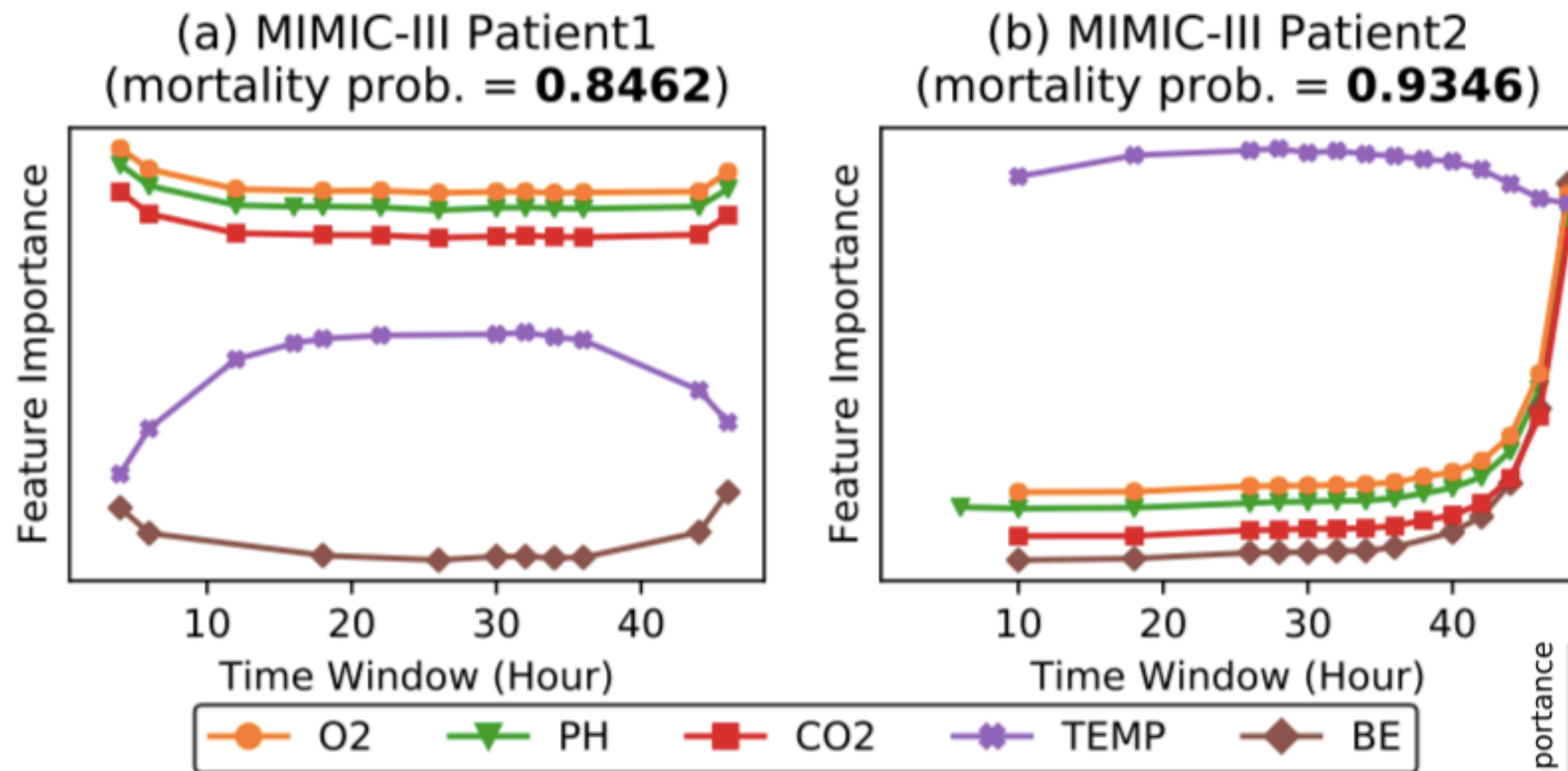TRACER: A Framework for Facilitating Accurate and Interpretable Analytics for High Stakes Applications, https://arxiv.org/abs/2003.12012
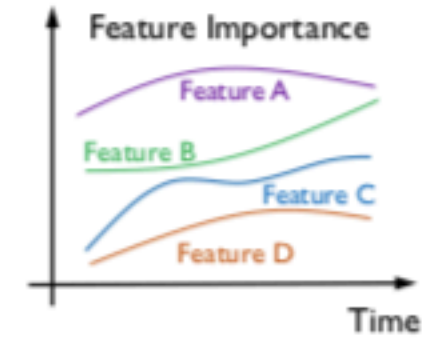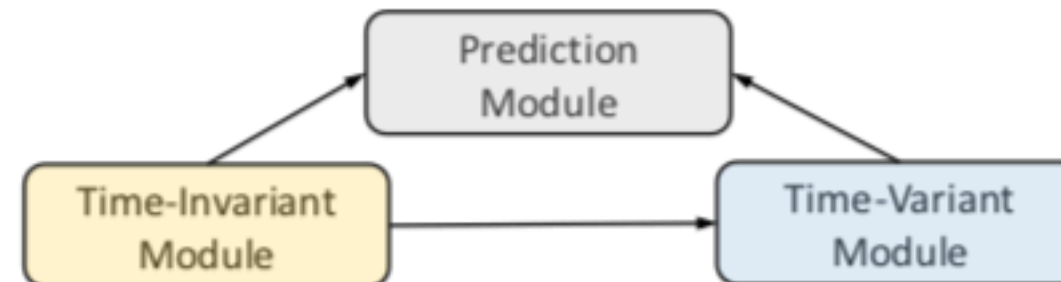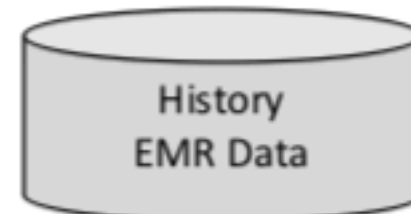
Patient-level:

TRACER pozwala zrozumieć dlaczego konkretni pacjenci umierają w szpitalu na przykładzie danych MIMIC III

(a) MIMIC-III Patient1 (mortality prob. = **0.8462**)

(b) MIMIC-III Patient2 (mortality prob. = **0.9346**)

Legend: O2, PH, CO2, TEMP, BE

Ilustracja jak zmienia się istotność zmiennych w czasie dla 4 zmiennych:
Serum Potassium (K),
Serum Sodium (NA),
Temperature (TEMP),
Mean Corpuscular Hemoglobin Concentration (MCHC)

(a) K (c) TEMP (b) NA (d) MCHC