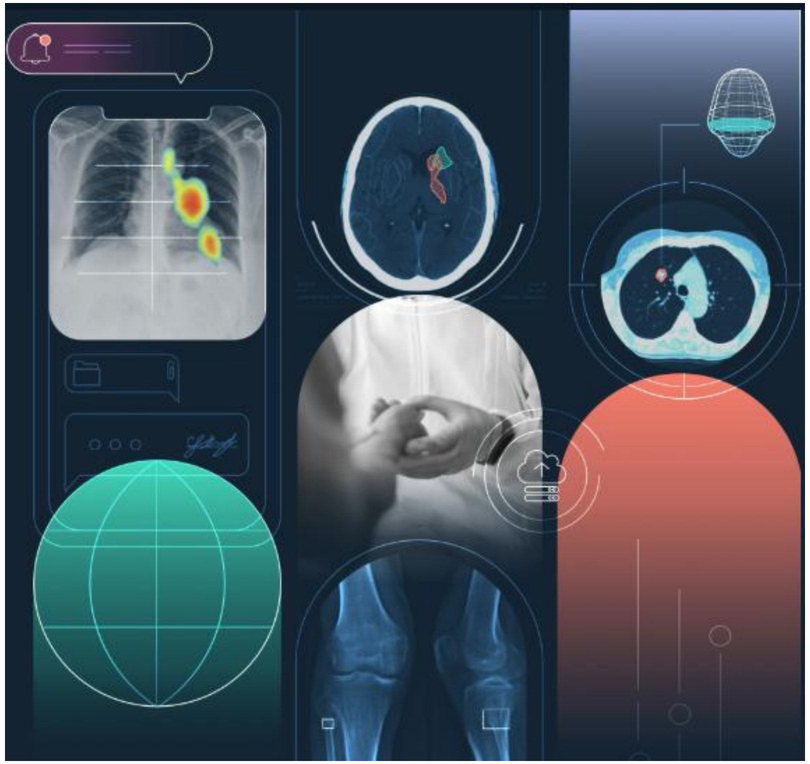
Case Study: AI in HealthCare

by

Rajdip Nayek

Assistant Professor,
Applied Mechanics Department,
IIT Delhi



AI in Healthcare – Opportunities, Challenge

What makes HealthCare Unique?

- **Healthcare impacts everyone**: Improving medical care leads to better quality of life
- **Variability in care:** Two patients with the same condition may receive **different treatments** depending on location, clinician experience, and available resources
- **Role of evidence**: Standardizing treatment through AI-driven evidence-based medicine can lead to more consistent outcomes

What does it mean to be healthy?

AI/ML can help define and predict health beyond traditional clinical definitions by incorporating genetic, behavioral, and environmental factors

Some stats from practicing physicians

35% of doctors report **burn-out.**[1]

56% **do not "have time"** to be **empathetic.**[2]

[1] Shanafelt, Tait D., et al. "Changes in burnout and satisfaction with work-life balance in physicians and the general US working population between 2011 and 2014." *Mayo Clinic Proceedings*. Vol. 90. No. 12. Elsevier, 2015.
[2] Riess, Helen, et al. "Empathy training for resident physicians: a randomized controlled trial of a neuroscience-informed curriculum." Journal of general internal medicine 27.10 (2012): 1280-1286

Why now?

- **Healthcare workforce crisis:** AI can help mitigate doctor shortages, particularly in rural areas

- **Explosion of healthcare data:** Electronic Health Records (EHRs), wearable devices, genomics, and imaging produce vast amounts of data

- **Advancements in AI & ML:** Deep learning (such as Idx-DR and EyeArt) has surpassed human-level performance in some medical tasks, such as diagnosing diabetic retinopathy

▸ Ophthalmol Sci. 2022 Sep 30;3(1):100228. doi: 10.1016/j.xops.2022.100228

Artificial Intelligence Detection of Diabetic Retinopathy

Subgroup Comparison of the EyeArt System with Ophthalmologists' Dilated Examinations

Jennifer Irene Lim [1,*], Carl D Regillo [2], SriniVas R Sadda [3], Eli Ipp [4], Malavika Bhaskaranand [5], Chaithanya Ramachandra [5], Kaushal Solanki [5], for the EyeArt Study Subgroup

▸ Author information ▸ Article notes ▸ Copyright and License information

PMCID: PMC9636573 PMID: 36345378

- **Computational power & accessibility:** Improved hardware and cloud computing enable scalable AI applications in medicine

Why AI is Crucial for India's Healthcare System?

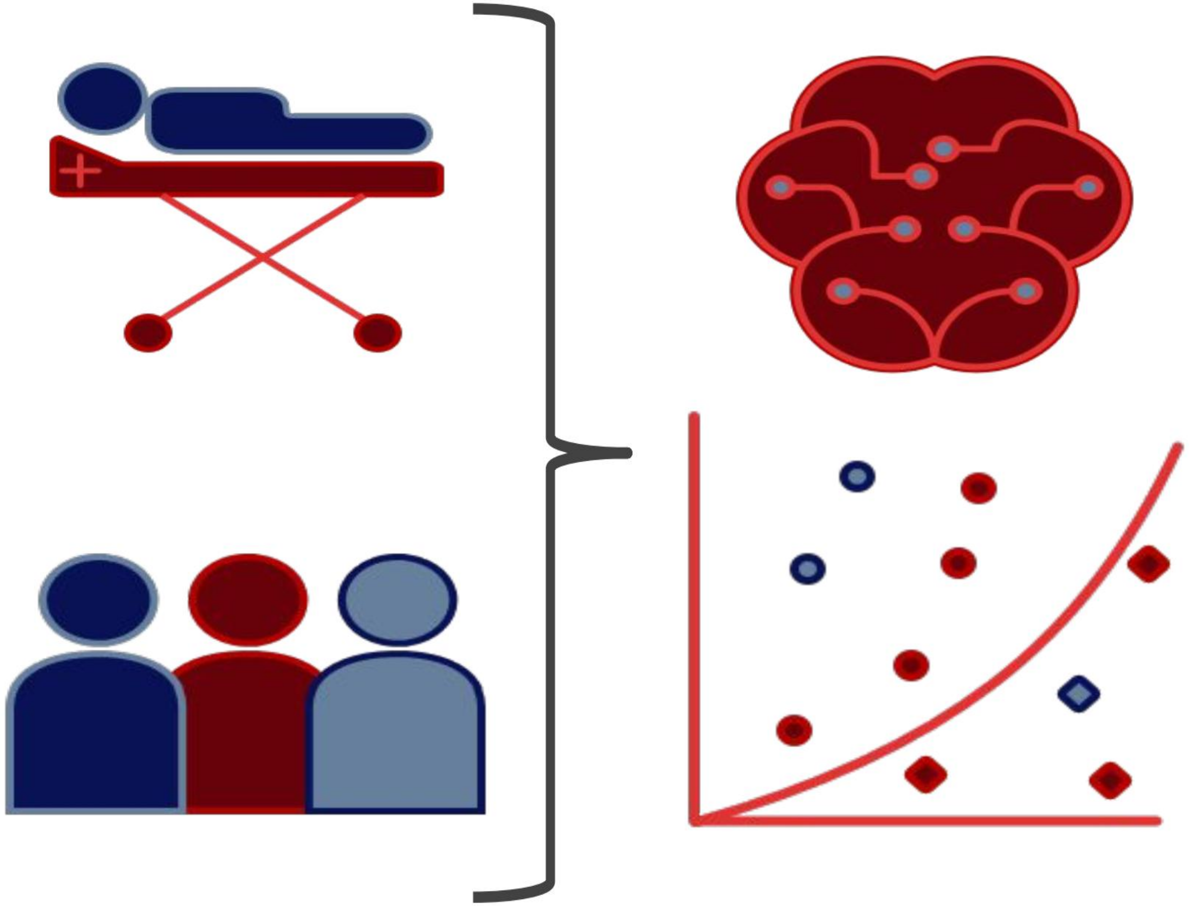
- **Doctor-Patient Ratio Challenge**
  - India has **1 doctor per 1,511 people** (WHO recommends 1:1000)
- **Healthcare Access Gap**
  - 70% of specialists are in urban areas; 65% of (Indian) people live in rural regions
- **Increasing Disease Burden**
  - Rise in **cardiovascular diseases, diabetes, and cancer**
- **Growing Healthcare Data**
  - Electronic Medical Records (EMRs), pathology reports, medical images

📊 **Statistic:** AI-driven early TB detection could reduce mortality by 20%

How AI/ML models learn in HealthCare?

Clinical data: from practicing doctors/hospitals

Medical knowledge: from books, RCTs, research papers

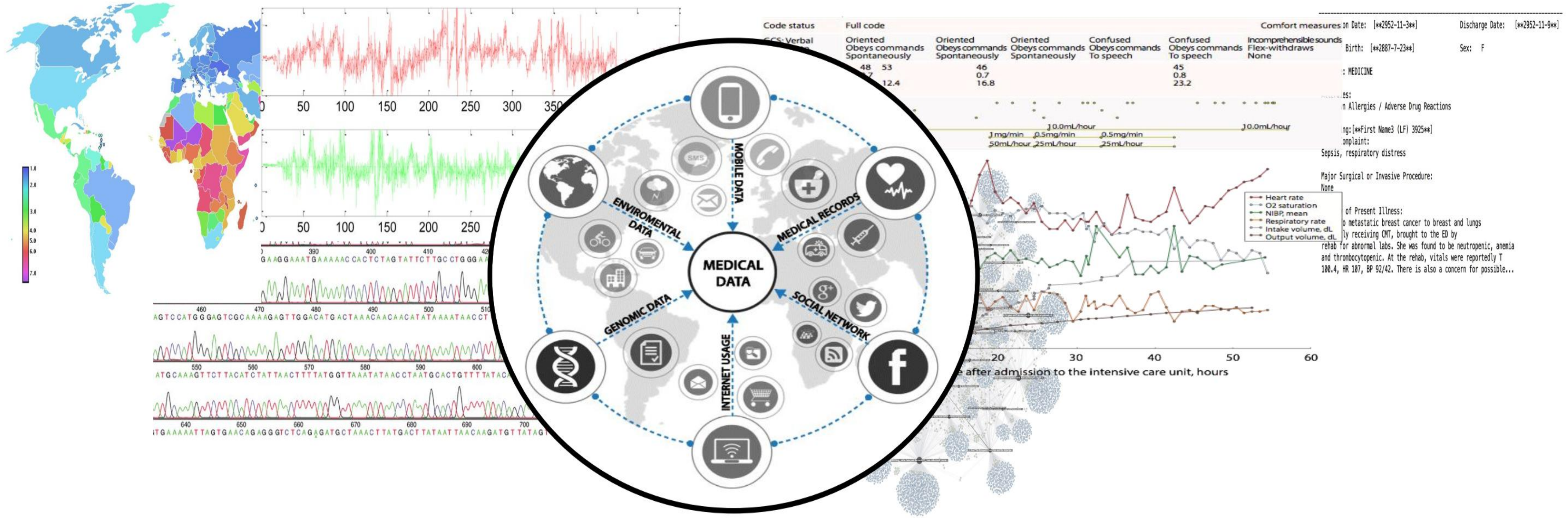


How AI works in Healthcare (Basic Mechanisms)?

- **Machine Learning:** AI models trained on medical data to recognize patterns (e.g., disease classification)
- **Deep Learning:** Neural networks used for complex tasks like image recognition (e.g., brain tumor detection in MRIs)
- **Natural Language Processing** (NLP): AI systems that analyze doctors' notes, clinical texts, and patient histories
- **Robotic Process Automation**: AI-powered chatbots, administrative automation, and robotic surgeries

Key Takeaway: AI does not replace doctors but **assists them in decision-making**

Data is (getting) Increasingly Available



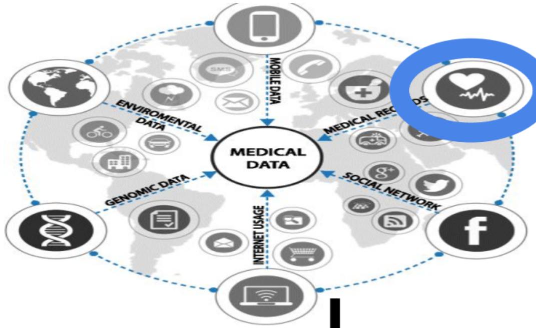
- **EHRs** (Electronic Health Records) are used by:
  - Over **80%** of UK hospitals
  - Over **60%** of Canadian practitioners
  - Around **40%** of Indian hospitals

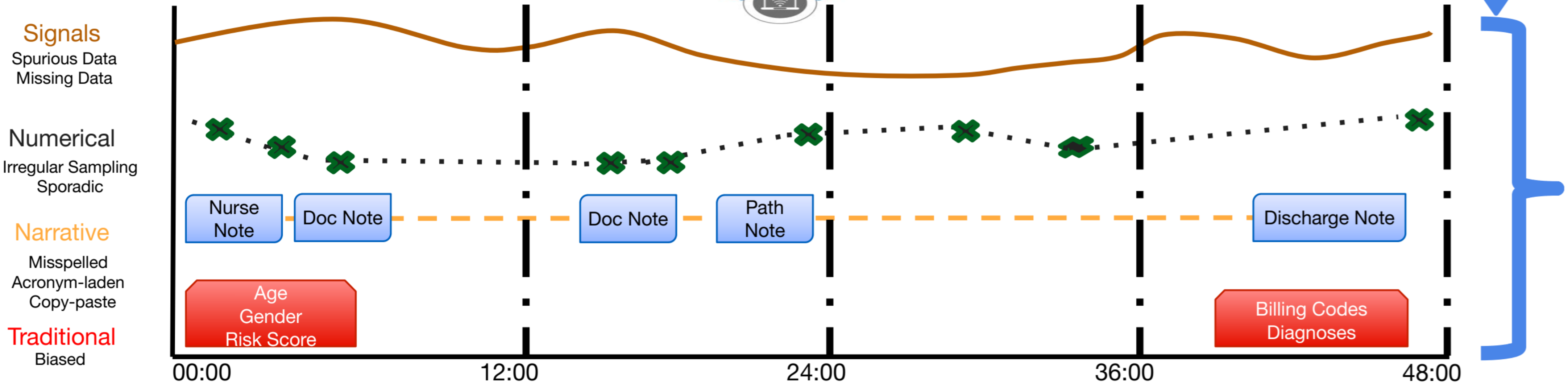
[1] *"Big Data in Health Care". The National Law Review. The Analysis Group, Inc.*
[2] Chang, Feng, and Nishi Gupta. "Progress in electronic medical record adoption in Canada." *Canadian Family Physician* 61.12 (2015): 1076-1084

Where to get the EHR?

- **MIMIC-III dataset** from Beth Israel Deaconess Medical Center ICU
- MIMIC is publicly available, and contains over 58,000 hospital admissions from approximately 38,600 adults

Multiparameter Intelligent Monitoring in Intensive Care (MIMIC)

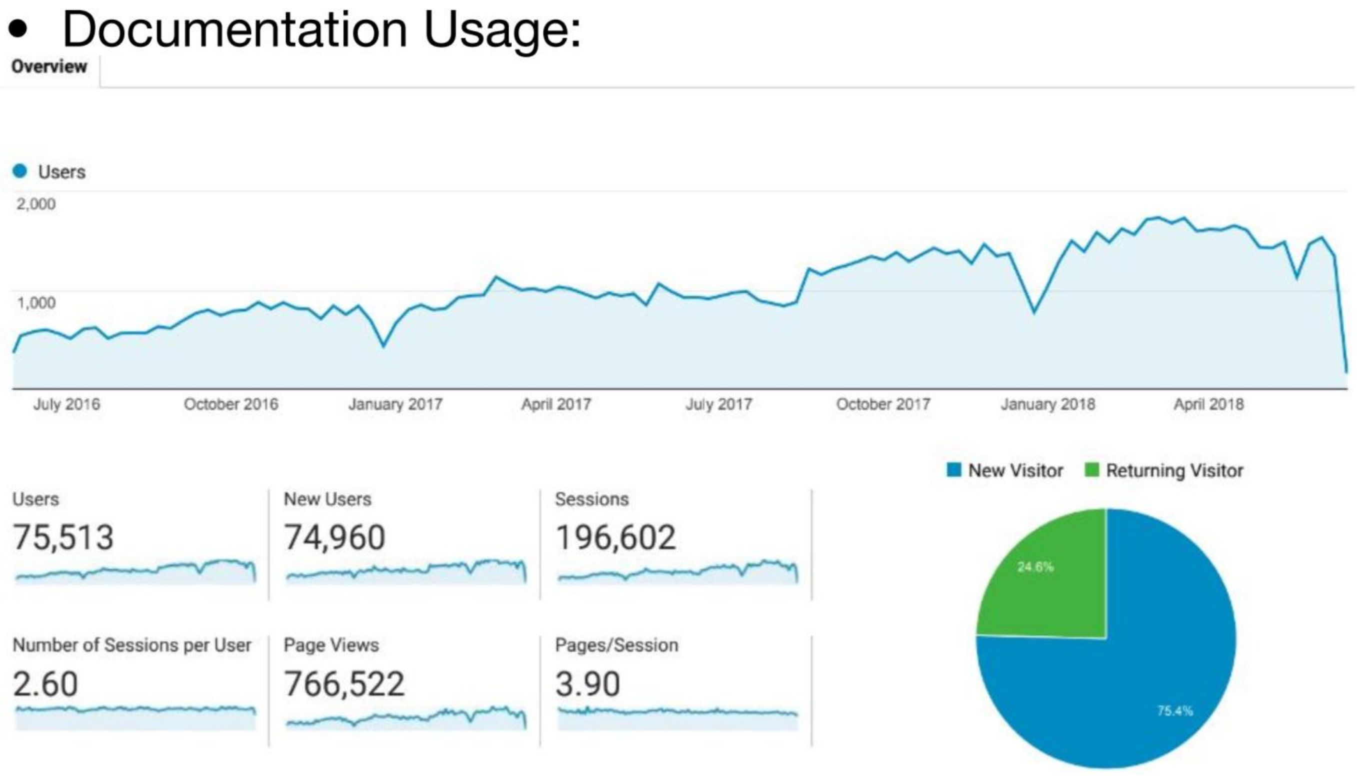




[1] Johnson, Alistair EW, et al. "MIMIC-III, a freely accessible critical care database." Scientific data 3 (2016)

MIMIC dataset is a huge resource

- **MIMIC-III dataset** from Beth Israel Deaconess Medical Center ICU



[1] Johnson, Alistair EW, et al. "MIMIC-III, a freely accessible critical care database." Scientific data 3 (2016)

AI/ML in HealthCare

DeepMind's new AI predicts kidney injury two days before it happens

New research from the Google-owned firm hints that AI may be a better way of assessing if someone is at risk of acute kidney injury. But there are still questions about how it handles patient data

The AI system was trained on over 620,000 distinct data points, with it eventually identifying **3,600 of them that were good predictors** of Acute Kidney Injury (AKI)

Only 6.32% of them were women, which meant that the AI system was less effective at predicting AKI when it was tested on **female patients**

An early sepsis prediction model utilizing machine learning and unbalanced data processing in a clinical context

Luyao Zhou [a], Min Shao [b], Cui Wang [b], Yu Wang [a]

The study included 2,385 patients, including 364 with sepsis, collected from the First Affiliated Hospital of Anhui Medical University and partner hospitals from April to July 2022

18 diagnostic features are used in the predictive model for early sepsis. The **Random Forest** model has the best performance among all the models, with an Area Under the Curve (AUC) of 87% and an F1-score of 77%. Moreover, the interpretation from the SHAP analysis is generally consistent with the current clinical situation

AI/ML in HealthCare

Automating artificial intelligence for medical decision-making

Model replaces the laborious process of annotating massive patient datasets by hand.

Rob Matheson | MIT News Office
August 5, 2019

The field of predictive analytics holds increasing promise for helping clinicians diagnose and treat patients. Machine-learning models can be trained to find patterns in patient data to aid in sepsis care, design safer chemotherapy regimens, and predict a patient's risk of having breast cancer or dying in the ICU, to name just a few examples.

Typically, training datasets consist of many sick and healthy subjects, but with relatively little data for each subject. Experts must then find just those aspects — or "features" — in the datasets that will be important for making predictions.

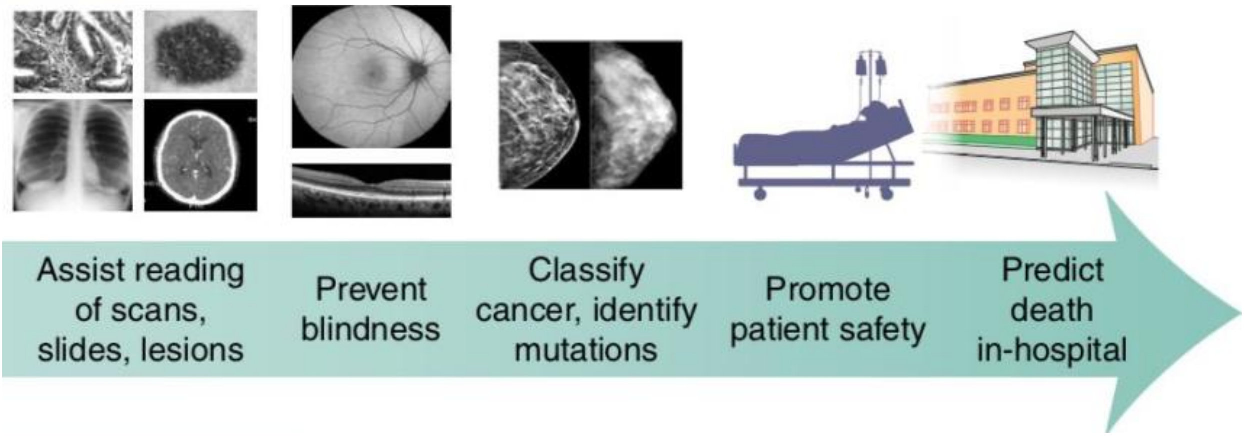
Amazon Web Services Teams with PHDA to Improve Care

8/7/2019

PITTSBURGH – In the latest sign of Pittsburgh's growing importance as a center of health care technology innovation, the Pittsburgh Health Data Alliance (PHDA) announced today that it is working closely with Amazon Web Services (AWS), an Amazon.com company, through a machine learning research sponsorship, to advance innovation in areas such as cancer diagnostics, precision medicine, voice-enabled technologies and medical imaging.

ML shows at/above Human-level Performance

Prediction	*n*	AUC
All-cause 3–12 month mortality	221,284	0.93*
Readmission	1,068	0.78
Sepsis	230,936	0.67
Septic shock	16,234	0.83
Severe sepsis	203,000	0.85@
Clostridium difficile infection	256,732	0.82++



Source: **High-performance medicine: the convergence of human and artificial intelligence** Eric Topol, Nature Medicine Jan 2019

Applications: Clinical Intervention Prediction

Real-time predictions of clinical interventions

Predicting what medical treatments a critically ill patient will need—**at the right time, while they are still in the ICU is very important**

Proceedings of Machine Learning for Healthcare 2017 JMLR W&C Track Volume 68

Clinical Intervention Prediction and Understanding with Deep Neural Networks

Harini Suresh [1] HSURESH@MIT.EDU

Nathan Hunt [1] NHUNT@MIT.EDU

Alistair Johnson [2] AEWJ@MIT.EDU

Leo Anthony Celi [2] LCELI@MIT.EDU

Peter Szolovits [1] PSZ@MIT.EDU

Marzyeh Ghassemi [1] MGHASSEM@MIT.EDU

[1] *Computer Science and Artificial Intelligence Lab, MIT, Cambridge, MA*
[2] *Laboratory for Computational Physiology, MIT, Cambridge, MA*

Applications: Clinical Intervention Prediction using ML

What data did they use?

- The dataset used is MIMIC-III
- It includes over 58,000 hospital admissions from 38,600 adult patients
- The study focuses on patients who:
  - are **15 years or older**
  - had **ICU stays lasting between 12 to 240 hours**
  - only their **first ICU stay** is considered to avoid duplicate cases
- After filtering, **34148 unique ICU stays** remain in the dataset

Applications: Clinical Intervention Prediction using ML

What data did they use?

For each ICU patient, three types of data are extracted:

1. **Static variables (5 features)**
   - These are fixed characteristics like **age, gender, and other patient demographics**
   - They do not change over time but are replicated across all time steps for consistency
2. **Time varying Vitals and Lab Results (29 features):**
   - These are physiological measurements such as **oxygen saturation, blood urea nitrogen, heart rate, etc.**
   - They are **timestamped and rounded to the nearest hour**
   - If multiple measurements occur within the same hour, their values are **averaged**
3. **De-identified Clinical Notes**
   - Text-based clinical notes written by doctors and nurses during the ICU stay
   - Extracted as a time-series and processed using **Latent Dirichlet Allocation (LDA)** to create structured topic-based representations

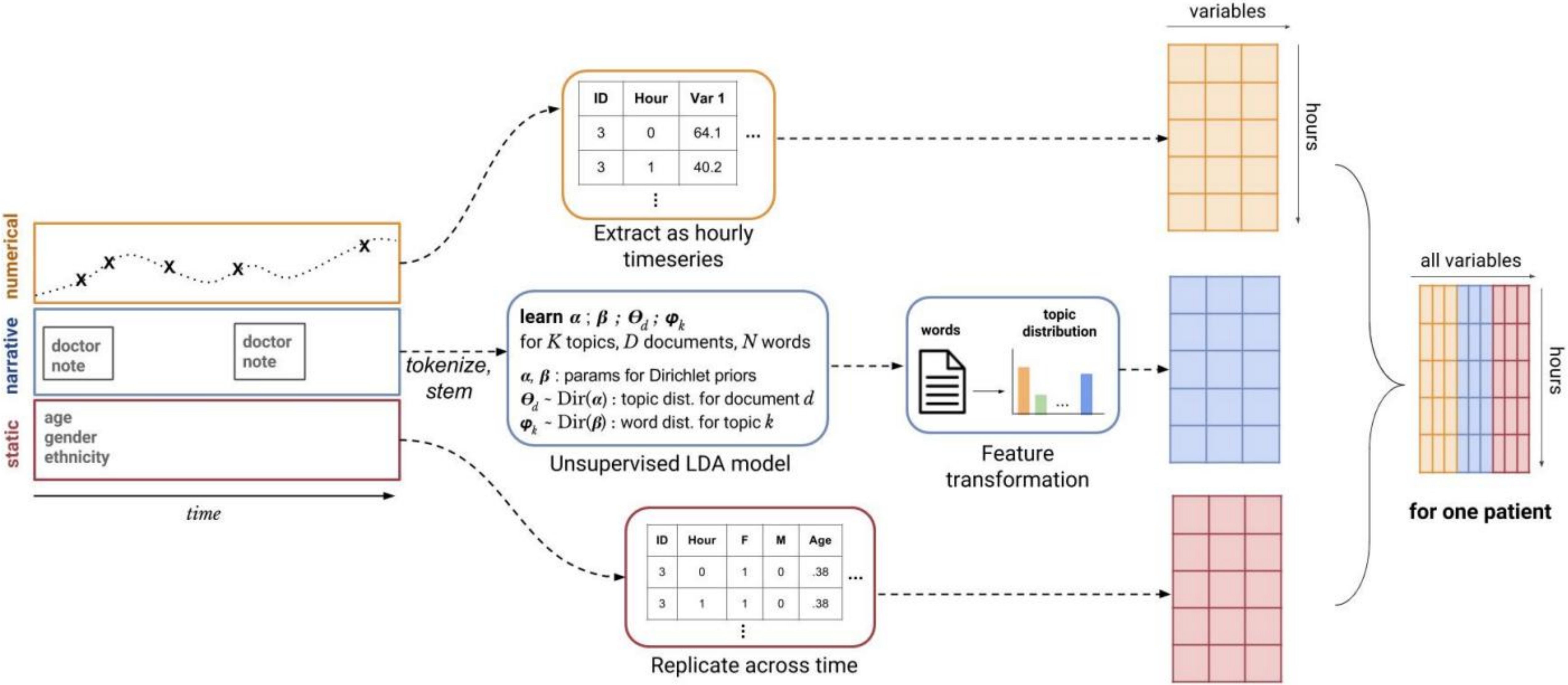
Applications: Clinical Intervention Prediction using ML

Representation of Notes and Vitals as Features for ML

1. **Clinical Notes Representation (Topic Modeling)**
   - Clinical notes are converted into **50-dimensional topic vectors** using **LDA**
   - These vectors represent distributions of medical topics (e.g., respiratory issues, infections)
   - Notes are **aggregated over time:**
     - If a **note A** is recorded at hour 3 and **note B** at hour 7, then:
       - Hours **3-6** retain A's topic distribution
       - From **hour 7 onward**, the aggregated distribution of **A and B** is used
2. **Physiological Data Processing:** Each vital sign value is converted into a z-score $z = \frac{(x - \text{mean})}{\text{std deviation}}$
3. **Intervention data Inclusion:** A binary variable is added to indicate whether a patient is receiving a specific intervention at each time of the day, where the hour of the day is an integer (0-23)
4. **Final Feature Vector Representation:** Each patient's data (static, vitals, and notes) along with intervention state and time of the day are included into a **single feature vector per patient**

Applications: Clinical Intervention Prediction using ML

Representation of Notes and Vitals as Features for ML

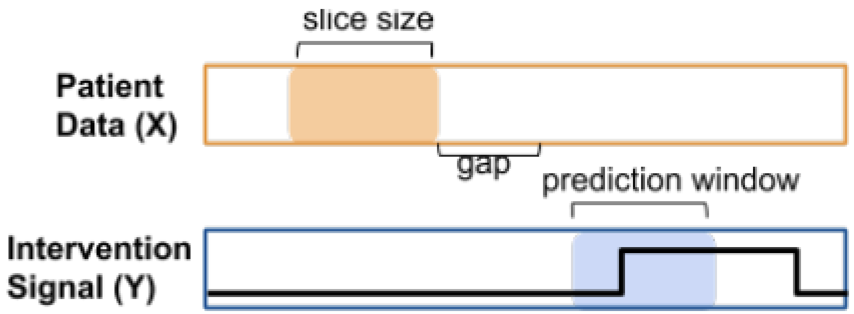


Applications: Clinical Intervention Prediction using ML

What kind of predictions are made?

Sliding window approach is used for time-series forecast

- Each patient record is split into **6-hour chunks** using a sliding window
- The model predicts interventions **4 hours ahead**, with a **6-hour gap** in between



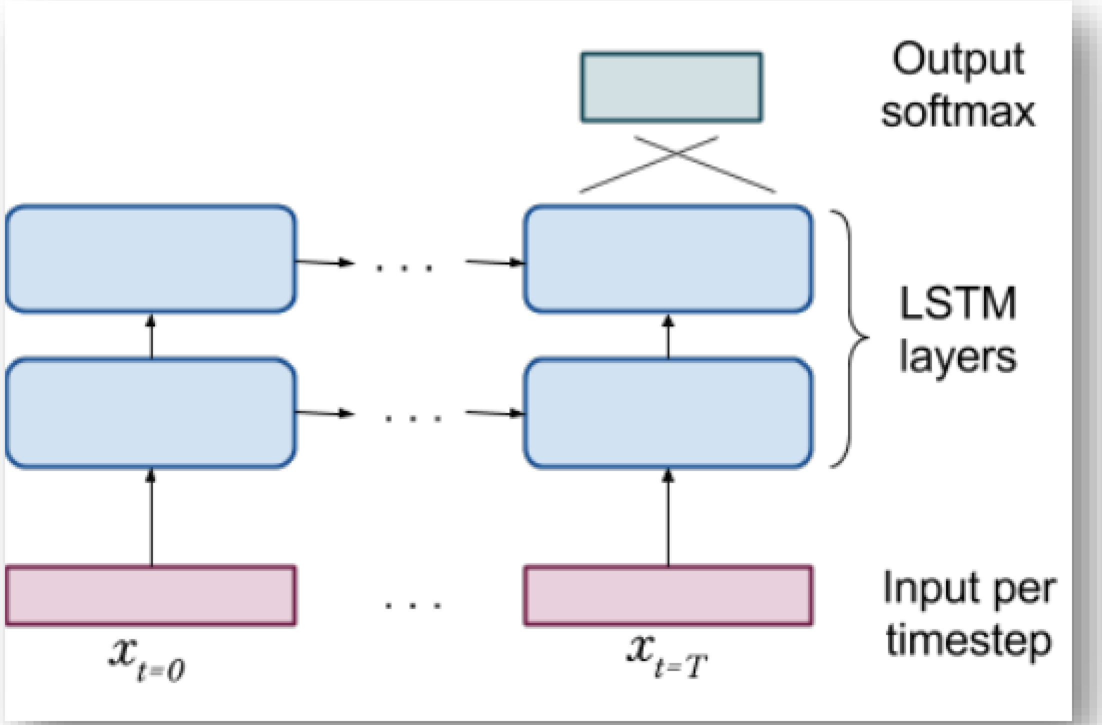
Prediction of Ventilation & Vasopressors (Four Categories):

- **Onset** – Patient transitions from **not receiving** intervention (0) to **receiving** it (1)
- **Wean** – Patient transitions from **receiving** intervention (1) to **not receiving** it (0)
- **Stay On** – Patient remains **on the intervention** for the entire window (1)
- **Stay Off** – Patient remains **off the intervention** for the entire window (0)

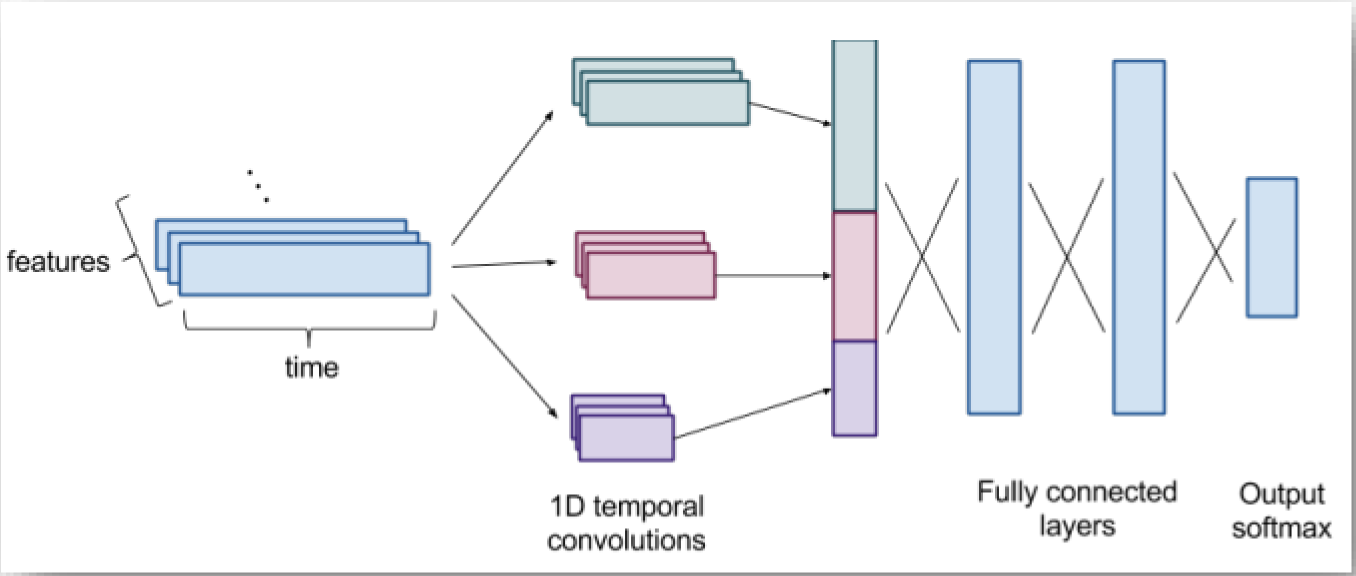
Applications: Clinical Intervention Prediction using ML

ML methods used?

Long-short-term-memory (LSTM) Network



Convolution Neural Network (CNN)



- Dataset is split into **70% training**, **10% validation**, and **20% test**
- **Stratified splitting** ensures class distribution remains balanced across all sets
- Training stops based on **AUC performance on the validation set** to prevent overfitting

Applications: Clinical Intervention Prediction using ML

Some results

Task	Model	VENT	NI-VENT	VASO	COL BOL	CRYS BOL
Onset AUC	Baseline	0.60	0.66	0.43	0.65	0.67
	LSTM	**0.75**	**0.76**	0.76	**0.72**	**0.71**
	CNN	0.62	0.73	**0.77**	0.70	0.69
Wean AUC	Baseline	0.83	0.71	0.74	-	-
	LSTM	0.90	**0.81**	**0.91**	-	-
	CNN	**0.91**	0.80	**0.91**	-	-
Stay On AUC	Baseline	0.50	0.79	0.55	-	-
	LSTM	**0.97**	**0.86**	0.95	-	-
	CNN	0.96	**0.86**	**0.96**	-	-
Stay Off AUC	Baseline	0.94	0.71	0.93	-	-
	LSTM	**0.97**	**0.86**	0.95	-	-
	CNN	0.95	**0.86**	**0.96**	-	-

Onset – Patient transitions from **not receiving** intervention (0) to **receiving** it (1)

Wean – Patient transitions from **receiving** intervention (1) to **not receiving** it (0)

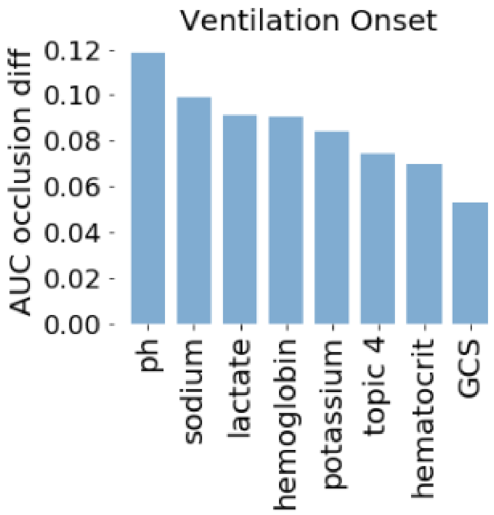
Stay On – Patient remains **on the intervention** for the entire window (1)

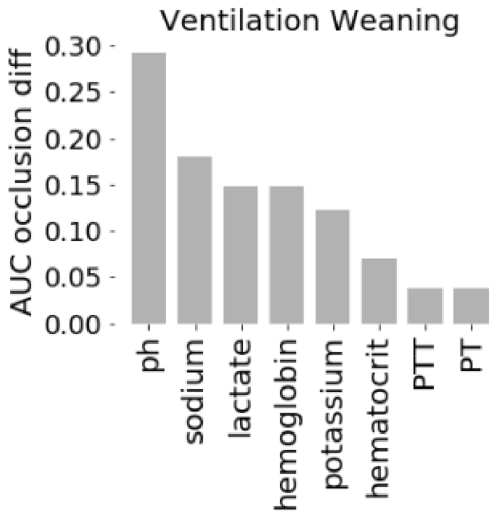
Stay Off – Patient remains **off the intervention** for the entire window (0)

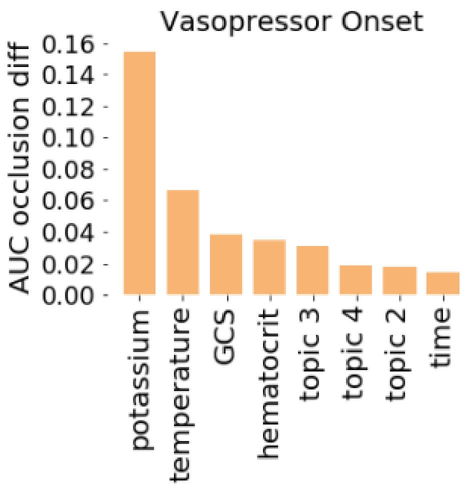
Baseline - Logistic regression (with L_2-regularization)

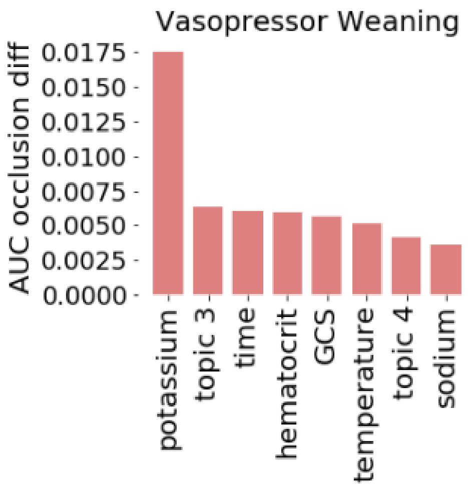
Applications: Clinical Intervention Prediction using ML

Occlusions used to understand which features were more important in prediction

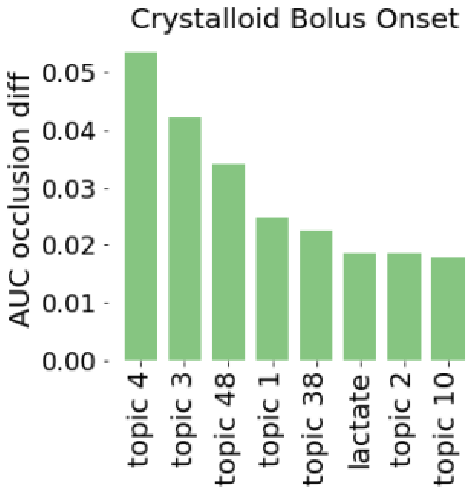


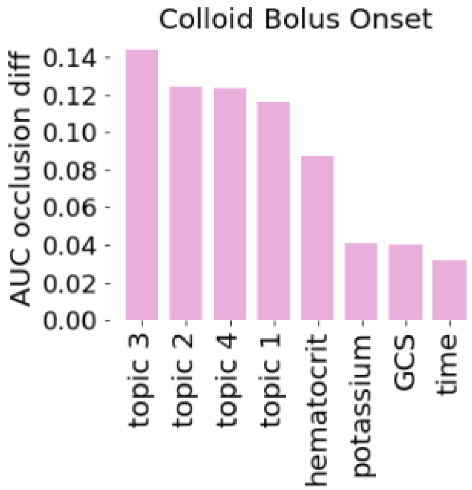






For **interpretability,** features are removed (**occluded**) one by one and replaced with noise, and the impact on predictions is analyzed to determine feature importance





Manifestation of bias in medicine

1. **Historical Bias in Medical Data**:
   - AI models learn from past healthcare records, which often reflect existing disparities (e.g.,underrepresentation of certain groups in clinical trials)
   - Example: A study showed that Black patients were less likely to receive pain medication due to biased historical data

2. **Algorithmic Bias in AI-Driven Decisions**:
   - If an AI model is trained mostly on data from urban hospitals, it may not work well for rural patients
   - Example: AI used for skin cancer detection was less accurate for darker skin tones because training data was primarily from lighter-skinned individuals

How should ML deal with bias in medical data?

How should ML deal with bias in medical data?

1. **Bias Detection & Auditing:**
    - Machine learning models must be tested for biases before deployment
    - Regular fairness audits can help detect disparities in AI-generated medical recommendations

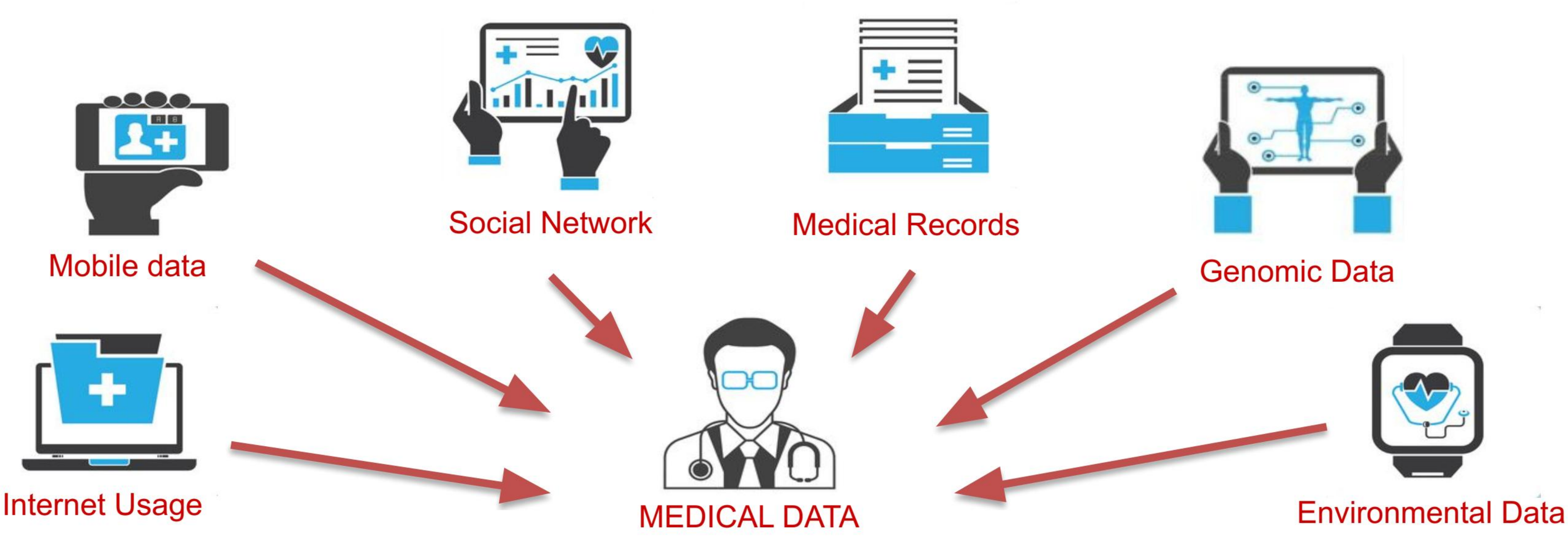
2. **Data Diversification**:
    - Ensure datasets represent a **diverse population** (gender, ethnicity, socioeconomic background) to avoid biased AI models
    - Example: Expanding genomics datasets to include **Indian genetic data** for personalized medicine

3. **Human-AI Collaboration:**
    - AI should **support doctors, not replace them**—ensuring decisions remain **context-aware and ethically sound**
    - AI recommendations should be explainable so that **doctors can override biased outputs** when necessary

4. **Regulatory & Ethical Oversight**:
    - Establishing **guidelines for AI fairness in healthcare** (e.g., India's AI ethics framework)
    - Ensuring transparency in AI-driven medical decisions through **explainability** techniques

Machine Learning for refining "What is healthy?"

Can we use **data** to **learn** what is **healthy** in a more **personalized** way?



Key AI Applications in Healthcare

- **Medical Imaging & Diagnosis** – AI detects diseases in **X-rays, MRIs, CT scans, and histopathology slides**
- **Predictive Analytics** – AI predicts disease outbreaks, patient deterioration, and treatment outcomes
- **Drug Discovery & Development** – AI accelerates drug testing (e.g., COVID-19 vaccines)
- **Robotic Surgery & Precision Medicine** – AI assists in minimally invasive surgery
- **Chatbots & Virtual Assistants** – AI-driven symptom checkers & appointment scheduling

Example: Google's AI model detects diabetic retinopathy in eye scans.

Promises and Potential Benefits

AI-Driven Diagnostics & Early Detection

- **AI-enhanced imaging**: Detecting tuberculosis, cancer, and diabetic retinopathy
- **Example: Google's DeepMind AI** detects diabetic retinopathy in India's rural clinics
- **Impact**: Faster, cost-effective screening where radiologists are scarce
- AI-driven **ECG interpretation** for heart disease detection (Apollo Hospitals uses AI to detect arrhythmias)