Conversion of DICOM ECG Images to Tabular Format for building Large Language Model in Diagnoses and Disease Progression of Cardiovascular Conditions

Bharath Potla, Dr Shivkumar J, Dr Sai Praveen Haranath, Dr Sujoy Kar, Dr Sangita Reddy

With the combined power of Apollo Hospitals, India's largest integrated healthcare system...

A 40-year legacy of transforming healthcare

73 Hospitals	**11,000+** Beds	**1570** Diagnostic Centres
200 Telemedicine Centres	**500** Clinics	**500+** Clinical Trials

Apollo 24|7

Largest omnichannel healthcare platform in India

700+ Collection Centres	**6,000+** Doctors	**12,500+** Pharma SKUs

We have delivered exceptional care over the past 4 decades...

Apollo HOSPITALS

200Mn Lives Impacted	**17Mn+** Tele-Consults	Joint Commission International **8** Accredited hospitals

Apollo 24|7

23Mn+ Registrations	**700K+** Daily Active Users	**35K+** Daily Rx orders

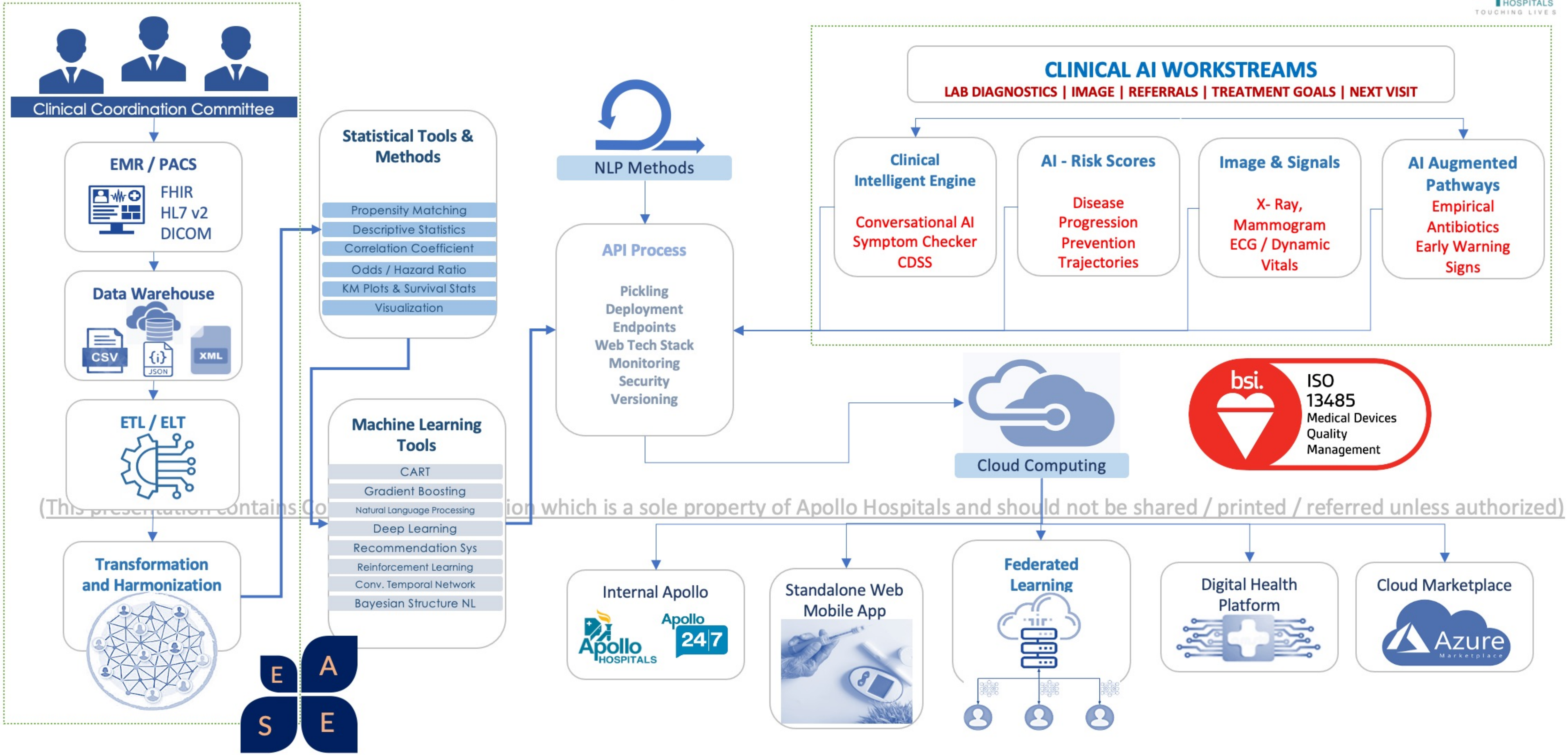
INSTITUTES OF TRANSPLANT Busiest Solid Organ Transplant Program in the world since 2012	24 hours EMERGENCY **500K+** Emergency calls served in 10+ years	Early innovator on E2E tech across the patient value chain
First combined **Elective Caesarean** and Robotic Assisted Radical Nephrectomy	**First and largest** Artificial Pulmonary Valve **implanted without surgery**	**First bone marrow transplant for Baby** with a novel mutation in blood

Among countless other firsts...

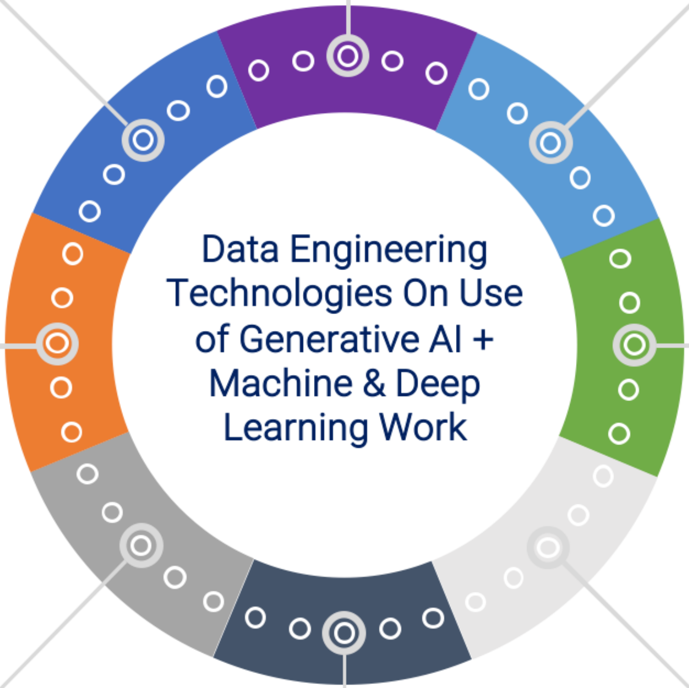
What will you do
today...

to transform the healthcare of
tomorrow?

Process Flow for Design – Development - Deployment



https://rdcu.be/c9MN3
https://doi.org/10.1007/s40012-023-00381-2



Entity Recognition

Identifying clinical entities in vast and diverse dataset

Language Translation

For Self-Care portion of discharge summaries

Entity Disambiguation

Associating the clinical entities with different codes and clinically relevant classification (medication + lab results)

Differentiated Database

Medications, Lab Logic, Self-Care, Triage, Pathways, Knowledge Base, HiPAR

Assertion

Identifying where in clinical data there are negation and over emphasizing a clinical term or decision

Question & Answers

Developing prompt methodologies to question about care directly to clinical database & get curated answers on diagnosis and treatment

Relation Extraction

Identifying relations between two clinical entities, identifying their correlations and merging different datasets to stich the context

Anonymization

Ensuring no Personal Identifiable Information and there in text of tabular formats of the data

DICOM Encapsulated ECGs

Importance of ECG Data
ECG (Electrocardiogram) is a fundamental diagnostic tool in Cardiology. It records the electrical activity of the heart and helps in diagnosing various heart conditions. In the digital age, the need to store, retrieve, and analyze ECG data efficiently is paramount. DICOM provides a standardized format for medical images, including ECGs, ensuring interoperability and data integrity.

Storing and Retrieving DICOM Encapsulated ECGs
DICOM allows for the encapsulation of ECG data, enabling healthcare organizations to store and manage ECGs alongside other medical imaging data. This standardized format ensures data consistency and simplifies data management.
•DICOM headers store patient information, acquisition details, and more.
•Encapsulated ECGs can be linked to patient records for easy retrieval.

Challenges and Opportunities While storing and retrieving DICOM encapsulated ECGs is essential, the real power lies in leveraging this data for improved patient care. Machine learning offers exciting opportunities for analysis and prediction based on ECG data.

HF Study

Objective

Apollo HOSPITALS
TOUCHING LIVES

#1 - Machine picks from an ECG
(As per Standard Communication Protocol) –

- HFrEF/HFpEF/HFmrEF (Abnormal) vs Normal ECG

(Propensity Matched)

- Differences in the wave / rhythm / rate patterns of HFrEF-HFpEF-HFmrEF vs Normal ECG

Add

#2 – Attributes for Heart Failure LLM

1) Echo - EF – >45% Ejection Fraction
2) Current Clinical Data - NYHA – Upgrade (Clinical Data) – Vitals + Comorbidities
3) Lab Marker – raised enzyme levels (pro BNP)
4) Medications
5) Revisit Longitudinal Data

Comparative Arms Only

HFpEF	HFrEF	HFmrEF	Normal

End Goal is for AI Model for Predicting HFpEF in next 5 years
Risk of HFpEF in 1, 2, 3 & 5 Years

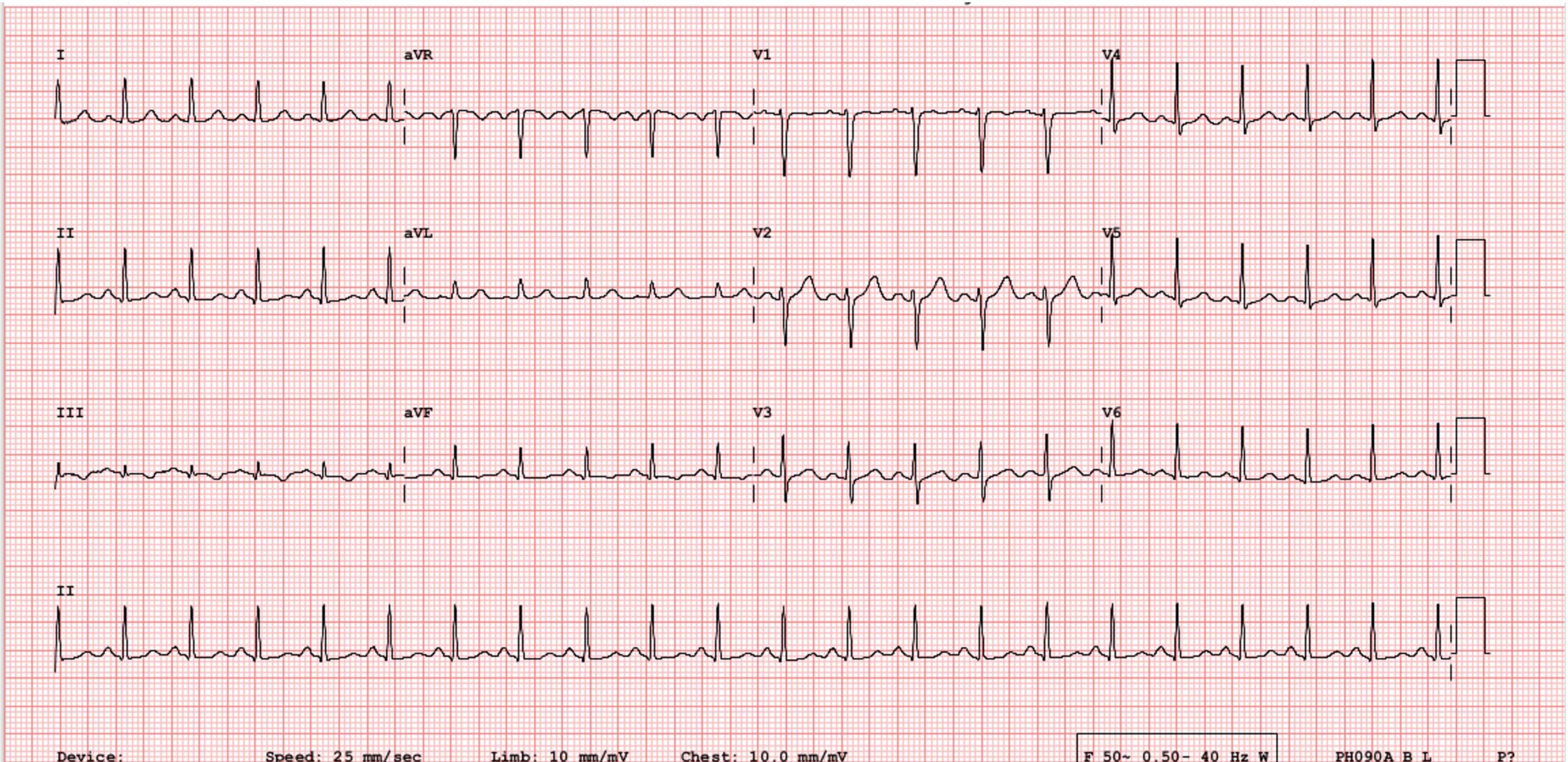
Clinical Decision Support
Lifestyle Modification
Heart Failure Registry

HF Study

ECG Conversion

Conversion of DICOM ECG Images to Tabular Format for building Large Language Model in Diagnoses and Disease Progression of Cardiovascular Conditions

Bharath Potla, Dr Shivkumar J, Dr Sai Praveen Haranath, Dr Sujoy Kar, Dr Sangita Reddy

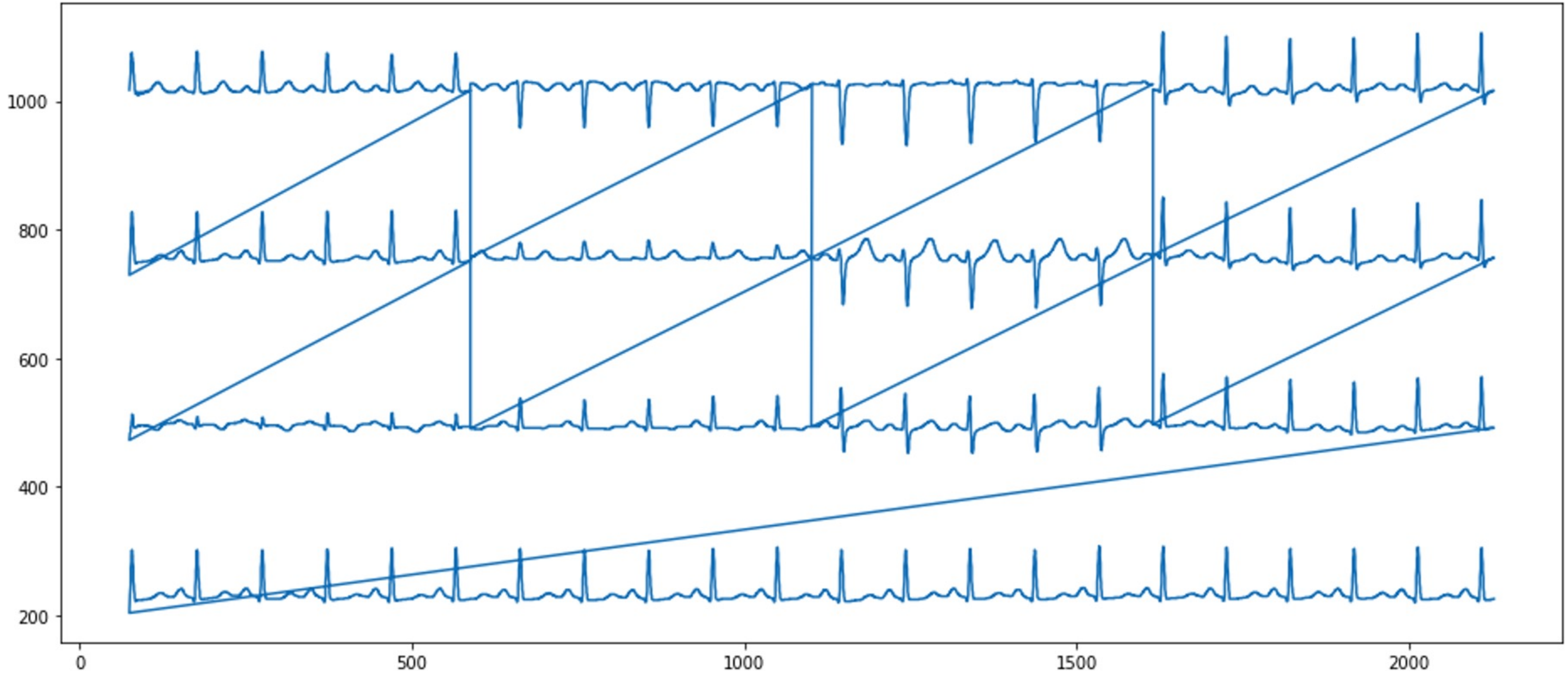


HF Study

ECG Conversion

Conversion of DICOM ECG Images to Tabular Format for building Large Language Model in Diagnoses and Disease Progression of Cardiovascular Conditions

Bharath Potla, Dr Shivkumar J, Dr Sai Praveen Haranath, Dr Sujoy Kar, Dr Sangita Reddy



We have been able to convert the whole ECG to near accuracy in tabular format. We have developed this as a home-grown API and converted at least 10K of Longitudinal Normal ECGs vs Abnormal ECGs (Heart Failure) – which is used to predict HF detection. The hazard model (time to event) model (predicting future heart failures) is in pipeline

HF Study
ECG Conversion – Predicting Accuracy in Categories

Conversion of DICOM ECG Images to Tabular Format for building Large Language Model in Diagnoses and Disease Progression of Cardiovascular Conditions

Bharath Potla, Dr Shivkumar J, Dr Sai Praveen Haranath, Dr Sujoy Kar, Dr Sangita Reddy

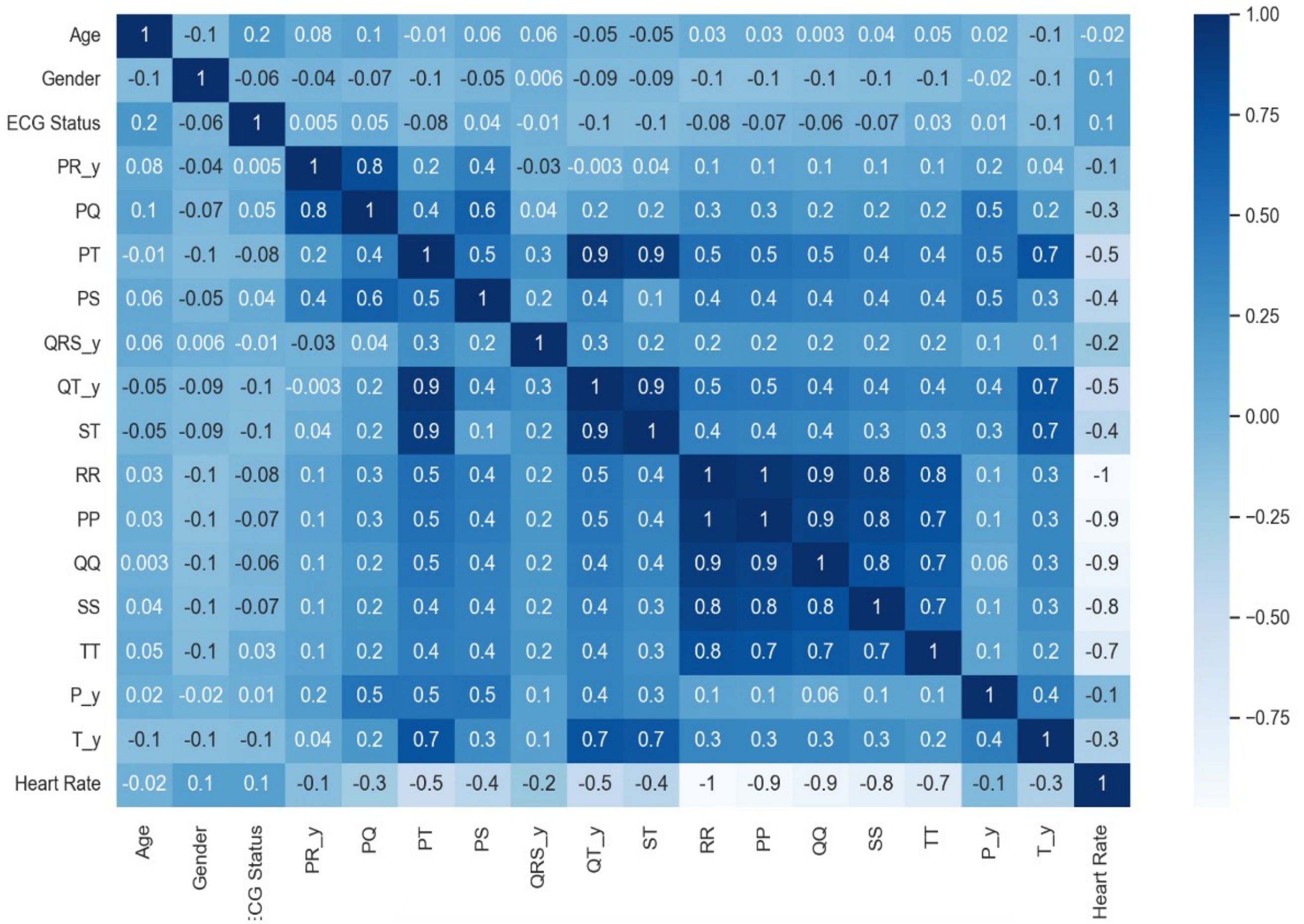
10K+| Preliminary XGB Classification Model on different Intervals in ECG

Abnormal records: 23%

Input Parameters : - X Parameters used: 'Age', 'Gender', 'PR', 'PQ', 'PT', 'PS', 'QRS', 'QT', 'ST', 'RR', 'PP', 'QQ', 'SS', 'TT', 'Heart Rate'.

Output Parameters : Y : Normal(0) & Abnormal(1)

Accuracy: 0.91 | ROC AUC: 0.86
Precision: 0.82 | Recall: 0.78 | F1 Score: 0.80

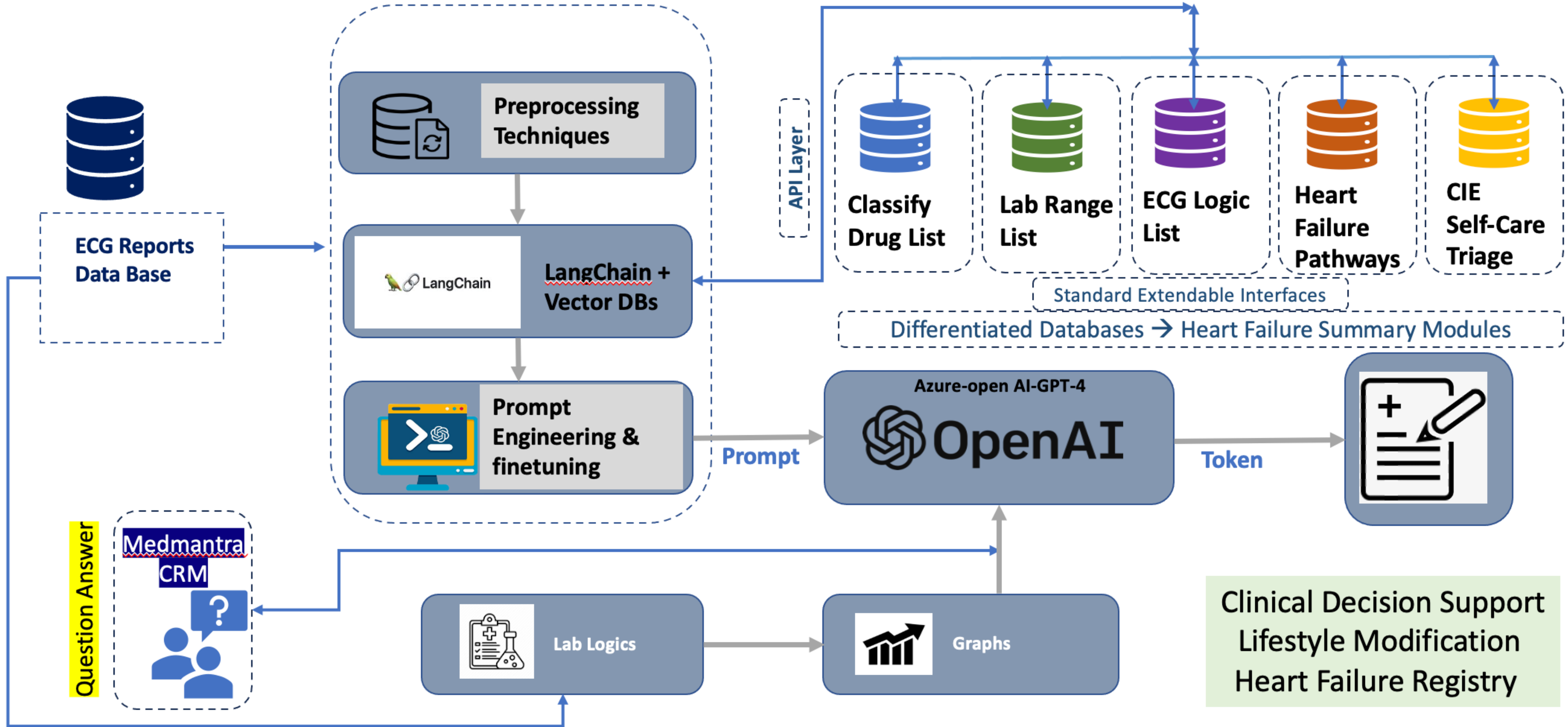


HF Study

ECG Conversion – Architecture for Unit ECG LLM Generation

Conversion of DICOM ECG Images to Tabular Format for building Large Language Model in Diagnoses and Disease Progression of Cardiovascular Conditions

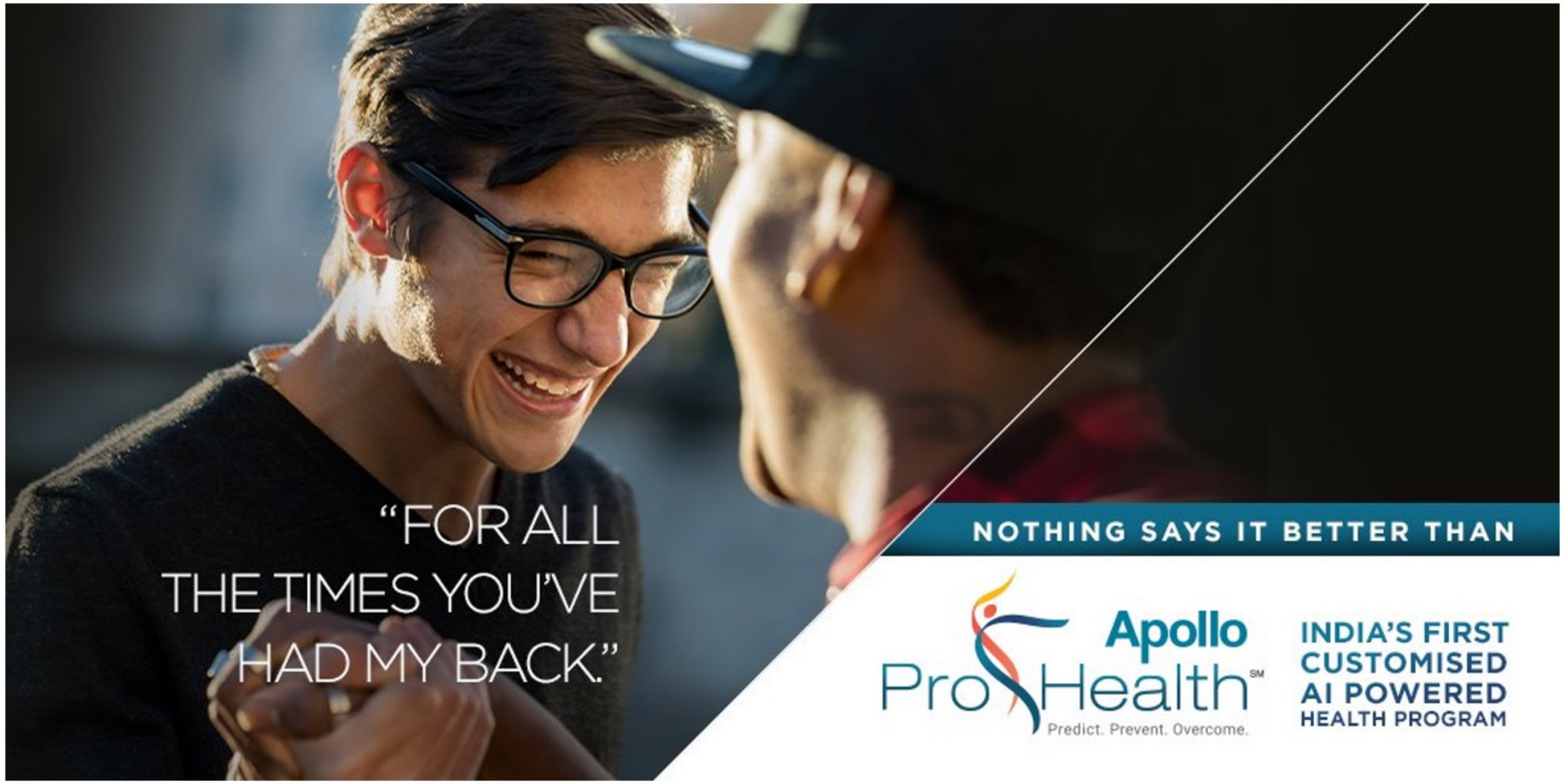
Bharath Potla, Dr Shivkumar J, Dr Sai Praveen Haranath, Dr Sujoy Kar, Dr Sangita Reddy

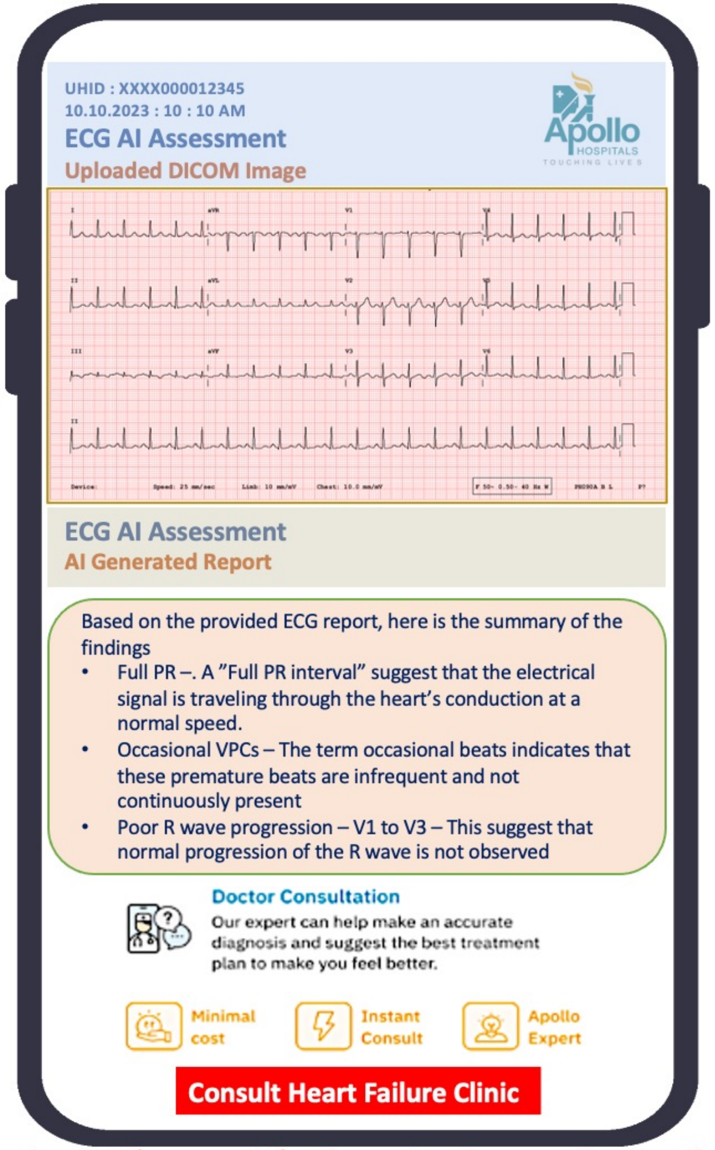


AI based Disease Progression & Risk Score Models

Deployment – Calibrate – Redesign - Redeploy

Value Capture: Deployment





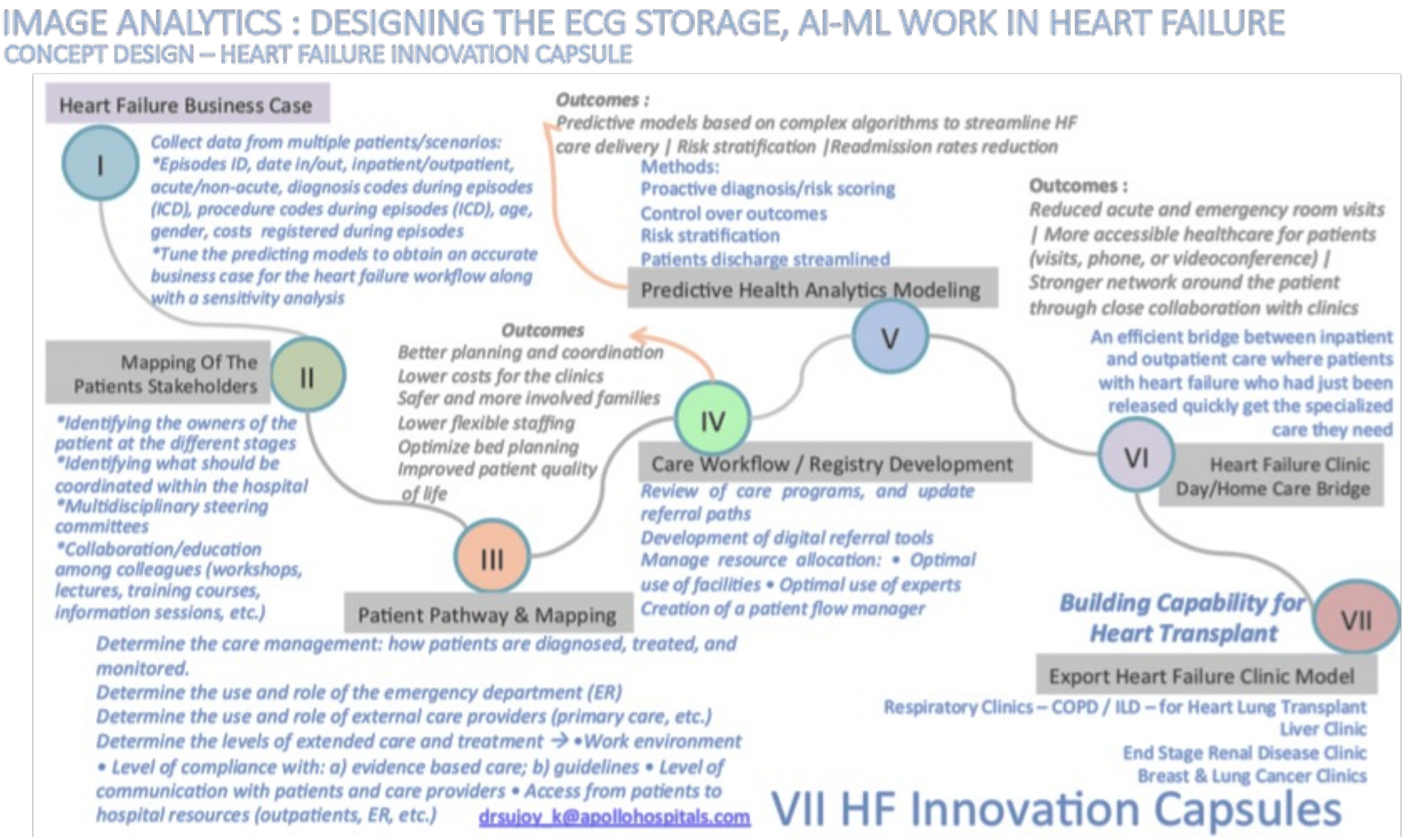
Apollo Throughput Optimisation (TOps Algorithms)

Level Up Clinical Protocols

	Pre-Anesthesia Algorithm	Early Warning Systems Wards / Telemetry	ER Triage to ICU	Discharge in 24/48 hours
Clinical Needs	• Risk Assessment tool for surgeries • Estimates surgical duration, blood loss and post operative patient placement	• Tool to help recognize early signs of clinical deterioration and trigger more intensive care • Prediction of Mortality | Risk Stratification | SHAP values | Advice for monitoring	• Identifies patient that could possibly transfer to ICU from ER • Risk of mortality in next 7 to 28 days	• Predicts probability of patient discharge in the next 24/48 hours • Use of Generative AI + Differentiated Database in building Discharge Summaries
Design & Development	• 347K Surgeries • 8 locations • 500+ surgery types over 18 months	• 145K Critical Patient (Anonymized) Data • Biphasic Model - Vitals + Clinical Features + Lab Data = XGBoost + Nested BERT	• Identifies patient that could possibly transfer to ICU from ER – Over 5K Data • Prototype Research - https://www.nature.com/articles/s41598-021-92146-7	• Collaborative Model with leading Organization • 160K Patient Data • Business Process Re-engineering
Ground Truth	Accuracy – **89%** + R^2 = **0.51**	Accuracy – **92%**	Accuracy - **93%**	Accuracy - **93%**
Impact	Improved OT Scheduling*	Remote Health Monitoring*	> 10K Risk Stratified (COVID)	ALOS Reduction* Improved Discharge Processes*

* Ongoing Prospective Use

Heart Failure Registry



How is the experience with Generative AI???

1

DIFFERENTIATED DATABASES

Generative AI in Healthcare is as good as curated content that you can build & provide as prompt

2

PROMPT BRITTLENESS

Variation in Prompt Syntax – change in wording, ordering, or selection of examples – make it unpredictable & unreliable

3

HALLUCINATION

Fluent and Natural generated texts which are unfaithful and / or undetermined.

Thanks

Any Questions?