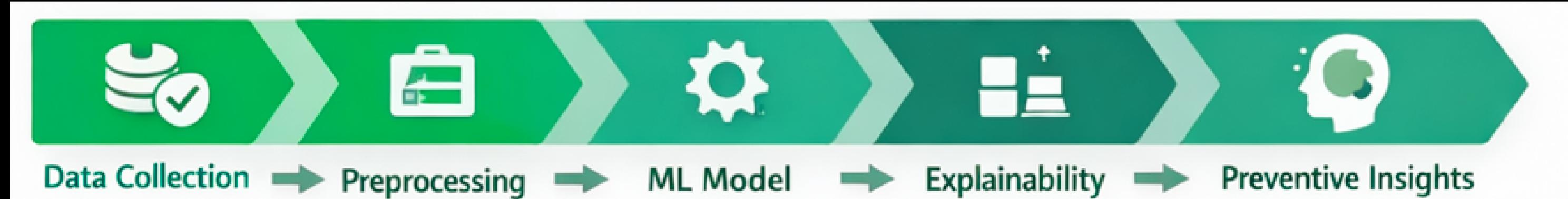


Our Roadmap:



Resources:

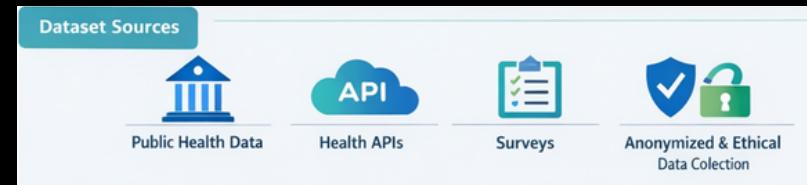
[NHANES \(Clinical Labs & Vitals\)](#)

[PhysioNet \(MIMIC-IV / eICU\)](#)

[CDC Open Data Portal \(Health & Lifestyle Data\)](#)

[WHO Global Health Observatory API](#)

[Columbia University: Disease Symptom Knowledge](#)



How was the data collected ?

- Collected through large-scale, government-led health surveys and clinical screening programs conducted by public health organizations (CDC, WHO).
- Data gathered using standardized medical examinations, laboratory tests, and structured questionnaires.
- Covers representative populations across regions and age groups for reliable risk analysis.
- Anonymized and ethically processed before being released via official open data portals and APIs

This data is highly relevant for early health risk prediction as it includes clinically validated biomarkers such as blood pressure, fasting glucose, HbA1c, cholesterol levels, and kidney filtration rates—key indicators of chronic diseases.

The data was collected through large-scale, government-led health surveys and screening programs, and made available via official public APIs and open data portals.

All datasets are anonymized, population-scale, and ethically collected, ensuring reliability, fairness, and privacy compliance. We have collected the datasets for 5 diseases now and we are focusing to enhance for them and add more if required.