Investigating Distance Decay and Social Determinants on Healthcare Outcomes in a Marginalized Patient Population

Anemone Kasasbeh, PhD
Lead Data Scientist

United Health Services Hospitals Inc.,

November 14th, 2023

Agenda

- Introduction
  - Healthcare Disparities
  - Marginalized Population
  - Explainable AI “Bridging the Gap between Application and Theory in Healthcare”
- Research Motivation
- Overall Problem Description and Severity
- AI application (Distance Decay and Social Determinants)

Healthcare Disparities

Refers to

Unequal Access to Healthcare

Driving Factors

Socio-economic Status (more than 50%)

Result in

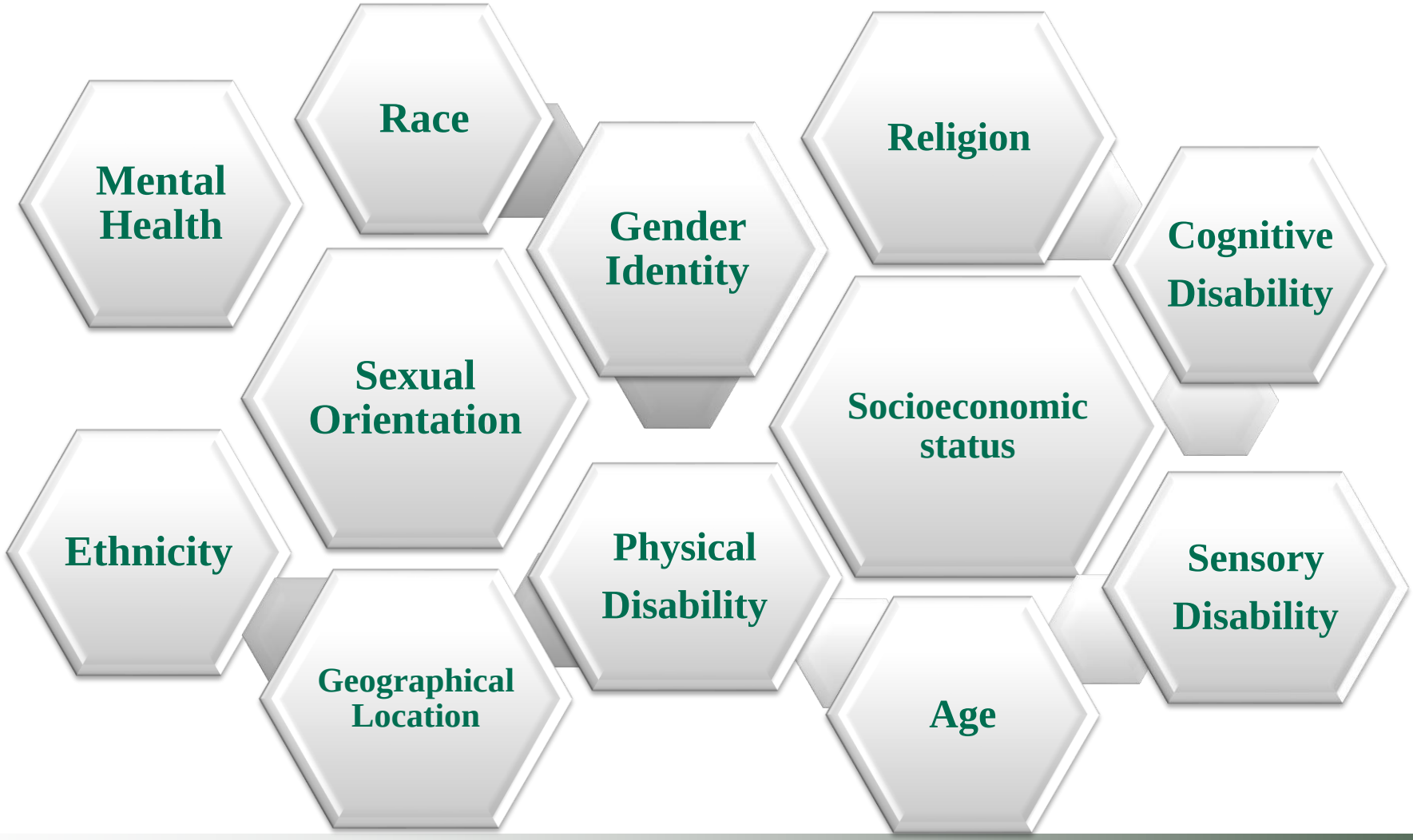
Poor Health Outcomes

Artificial Intelligence

Systematic Exclusion

Marginalized Patient Populations

According to the U.S. Department of Health :



Research Gap

- Systematic Exclusions in Emerging AI Applications while Limiting Marginalized Patient Population
  - Lack of Research
    - Google Scholar (Marginalized Population, Healthcare, Predictive Modeling) < 45

Marginalization Facts and AI ethical dilemmas

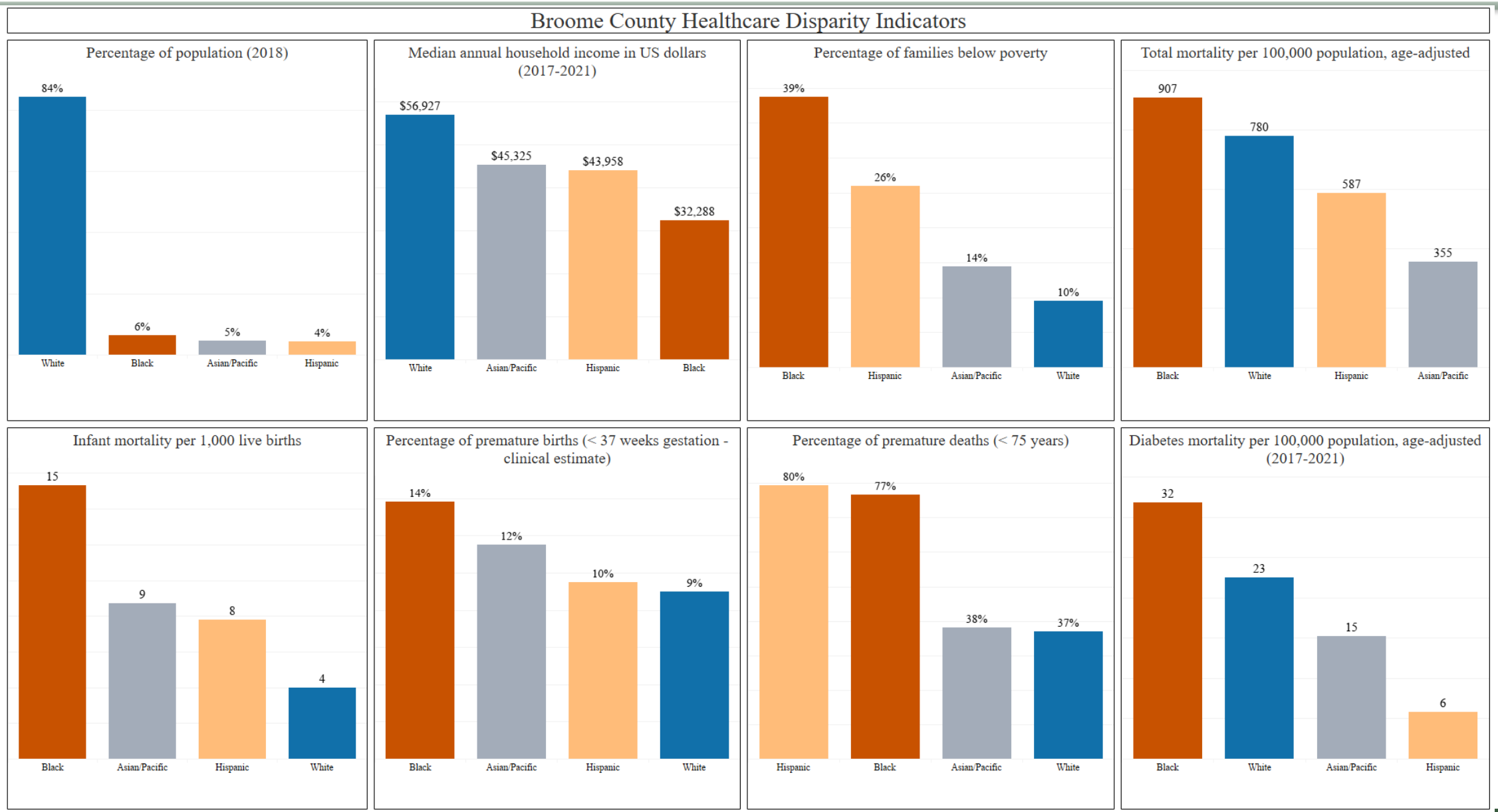
According to the CDC

- African Americans are 30% more likely to die from heart disease than non-Hispanic whites
- Hispanic Americans are 50% more likely to die from diabetes than non-Hispanic whites.
- African Americans are 60% more likely to have diabetes than non-Hispanic white Americans
- Low-Income Individuals are more likely to have uncontrolled high blood pressure
- Individuals with Disabilities have higher rates of chronic diseases

Marginalized Populations in this Research

Patients Residing in Broome County, NY

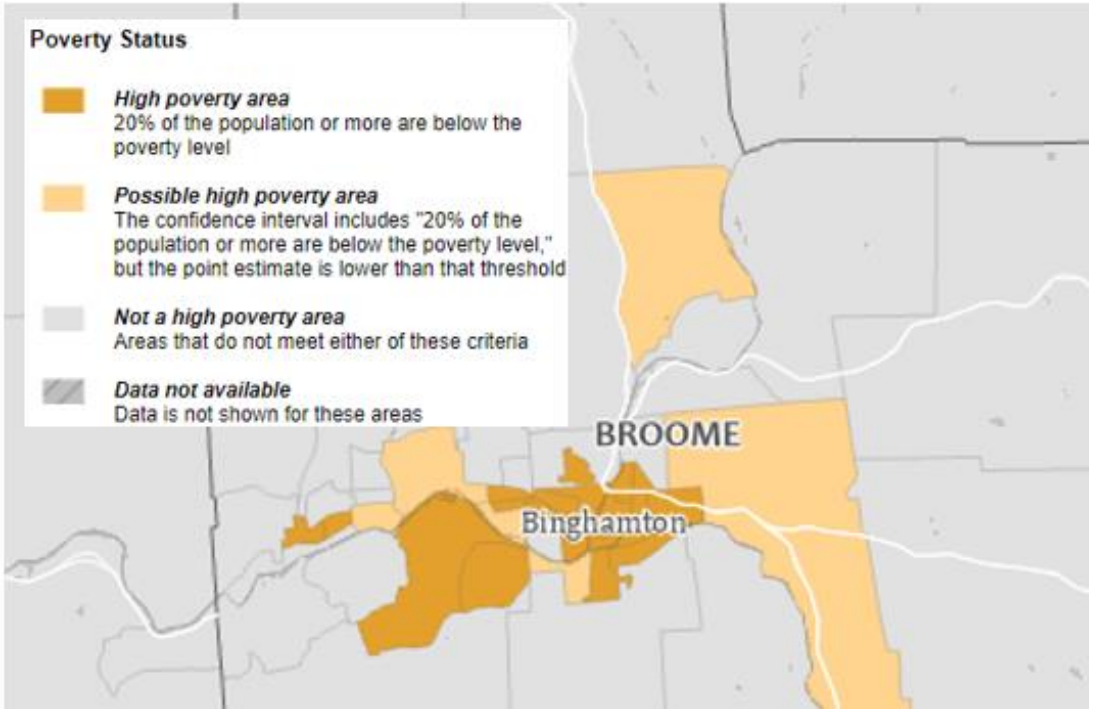
Healthcare Disparities Indicators at Broome County



This dashboard's data was obtained from Department of Health

Broome County and Marginalization

- Part of NY State but drastically different than NYC demographics.
- Poverty is 14.1% of the population in 2021
  - Higher than the national poverty rate.
- Percentage of persons 65 years and above is 20% of the population
  - Higher than the national rate.

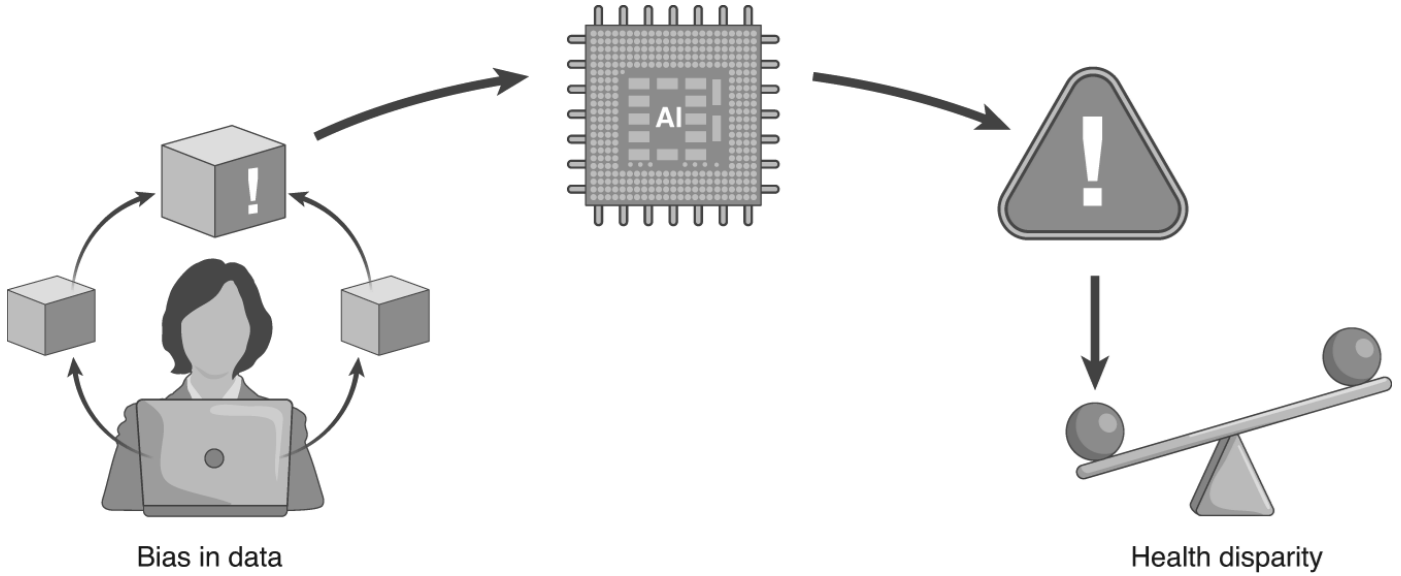


Poverty Status Viewer: With a Focus Area on Broome County, NY.
Source: U.S. Census Bureau (2021)

Research Motivation

Demonstrate, Highlight, Propose and Improve…

- ❑ Healthcare Disparities
- ❑ Inequity and Inequality
- ❑ AI Biasness
- ❑ Data Biasness
- ❑ Systemic Exclusions in Healthcare



Summary of Developed Application

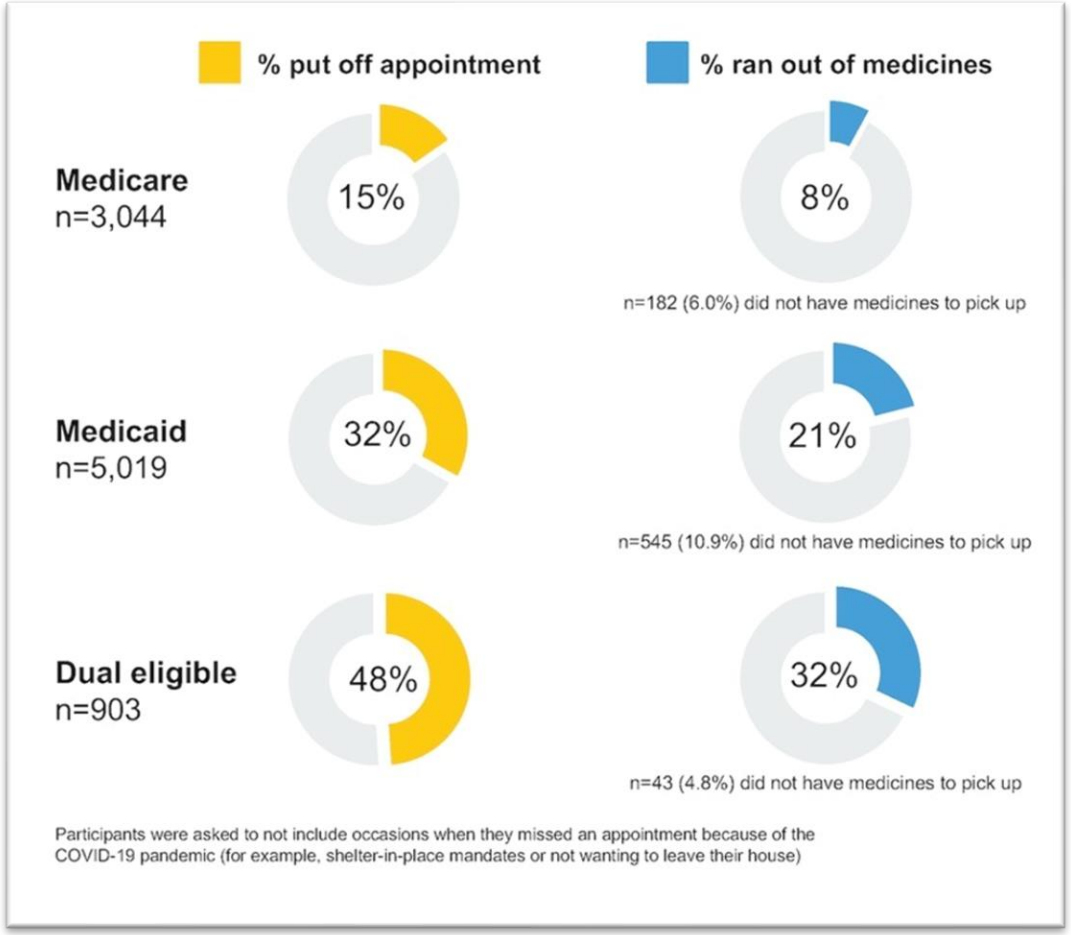
Application	Goal	Marginalized Population	Methodology	Integrated System
Investigating Distance Decay and Social Determinants on Healthcare Outcomes	Improve Access to Preventative Care	Broome County, NY	Machine learning (ML) Meta-Ensembles and Hybrid Models	From Data Acquisition to Automation (Patient List)

Study Goal/Objectives/General Approach

Goal	Objective	Approach
• Improve Marginalized Patient Population Healthcare Outcomes	• Recommend System Integrated Automated Frameworks to Provide Insights for Stakeholders	• Data Integration, Normalization • Data Preprocessing • Training, Validating, Testing • Predictions Generator • User Friendly Interface

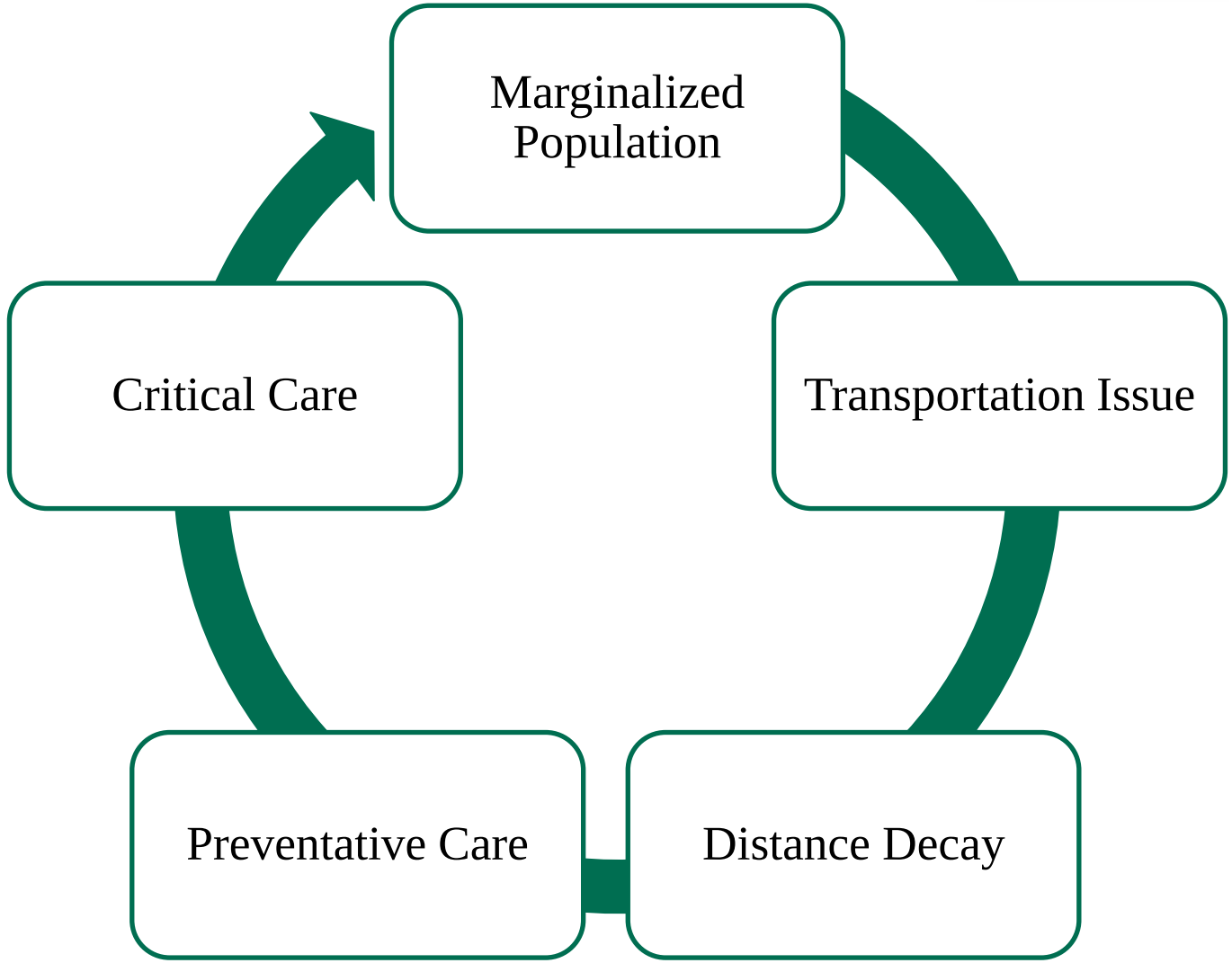
Introduction

- Substantial Transformations in Transportation and Healthcare Systems Using AI
- Distance Decay and Utilization of Healthcare
- Social Determinants and Spatial Interactions in Healthcare



Background and Problem Description

- Availability of transportation is key in conceptualizing ***Marginalized Patients*** access
- Longer Travel Distance and Transportation Availability are Linked to Healthcare Outcome
- Social Determinants and Economic Status



Research Gap

- Previous studies mainly relied on Traditional Statistics
- Machine Learning is crucial to Evaluate Confounding Factors in such Complex System
- No research in machine learning applications studied the

"Impact of Distance Decay on Healthcare Outcomes among Marginalized Patient Population"

- The first study to use meta-ensembles to study the Distance Decay in ***Marginalized Patients***

Motivation

- Improve Healthcare Outcomes in Marginalized Populations using AI-Based Integrated Systems

- Address/Highlight Social Determinants Association to Healthcare Outcomes
  - Awareness
  - Education

- Propose Early Intervention Workflow Practices

Model Objective/Input/Output

- Objective: Improve Office Utilization and Reduce Critical Care Utilization
- Inputs: 11 Features (Patient Related and Clinic Related)
- Output: Critical Care Group (High/Low)

Data Description and Features

- Hospital database.
- January 2021 and July 2022.
- Each row in the dataset represents a patient
- 43,152 unique patients, 38,812 no-shows, 394,116 attended visits, 3,923 emergency department (ED) visits, and 8,877 hospitalization episodes.
- Critical Care Group.

Factor Name	Inclusion	Type	Levels
Last Visit Date	Removed	Date/Time	-
Patient Primary Care Department	Kept	Categorical	8 levels (Primary Care Offices)
Patient Age	Removed	Categorical	-
Known Transportation Issue	Kept	Categorical	2 levels
Financial Class	Kept	Categorical	9 Levels (Medicaid, Medicare…etc.)
Distance between the patient's address and clinic	Kept	Categorical	3 Levels: (A, B, C)
Patient Generation	Kept	Categorical	6 levels: (Baby Boomer, Generation X, Generation Z, Greatest Generation, Millennials, Silent Generation)
Patient Sex	Kept	Categorical	2 levels
Patient Race	Kept	Categorical	17 levels
Missed Appointments	Transformed	Numerical	-
No-show Rate	Created	Categorical	2 levels (High, Low)
ED Visits	Transformed	Numerical	-
Hospitalizations	Transformed	Numerical	-
Critical Care group	Created	Categorical	2. levels (High, Low)

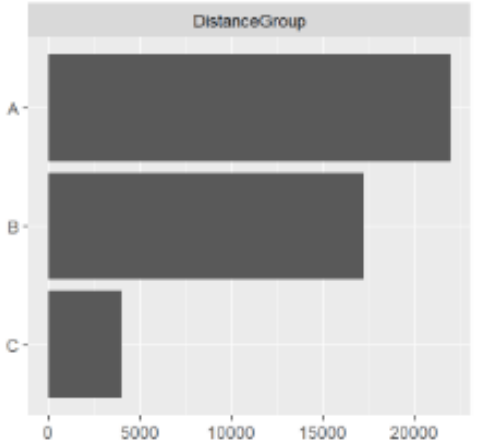
Data Exploration

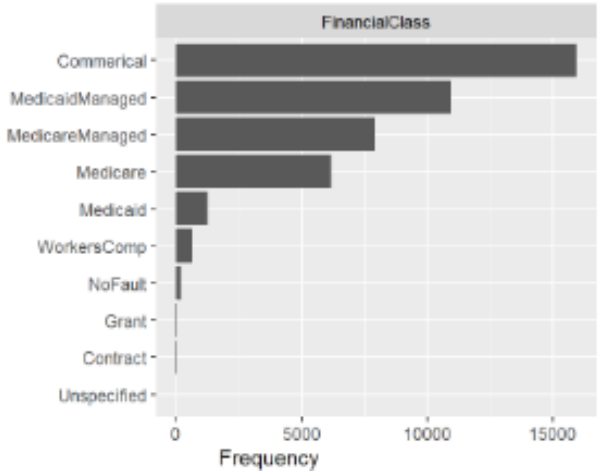
Measures of Healthcare System Utilization with Respect to Transportation

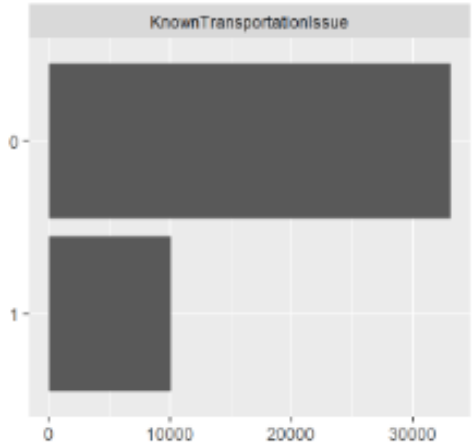
Known Transportation Issue	Average Office Visits Utilization	Average Critical Care Utilization
No	91.5%	2.5%
Yes	79%	6.5%

Measures of Healthcare System Delivery with Respect to Race and Transportation

Transportation Issue	Race	Average Office Visits Utilization	Average Critical Care Utilization
No	Black or African American	81%	3.6%
	White	93%	2.5%
Yes	Black or African American	69%	7.5%
	White	81%	6.5%



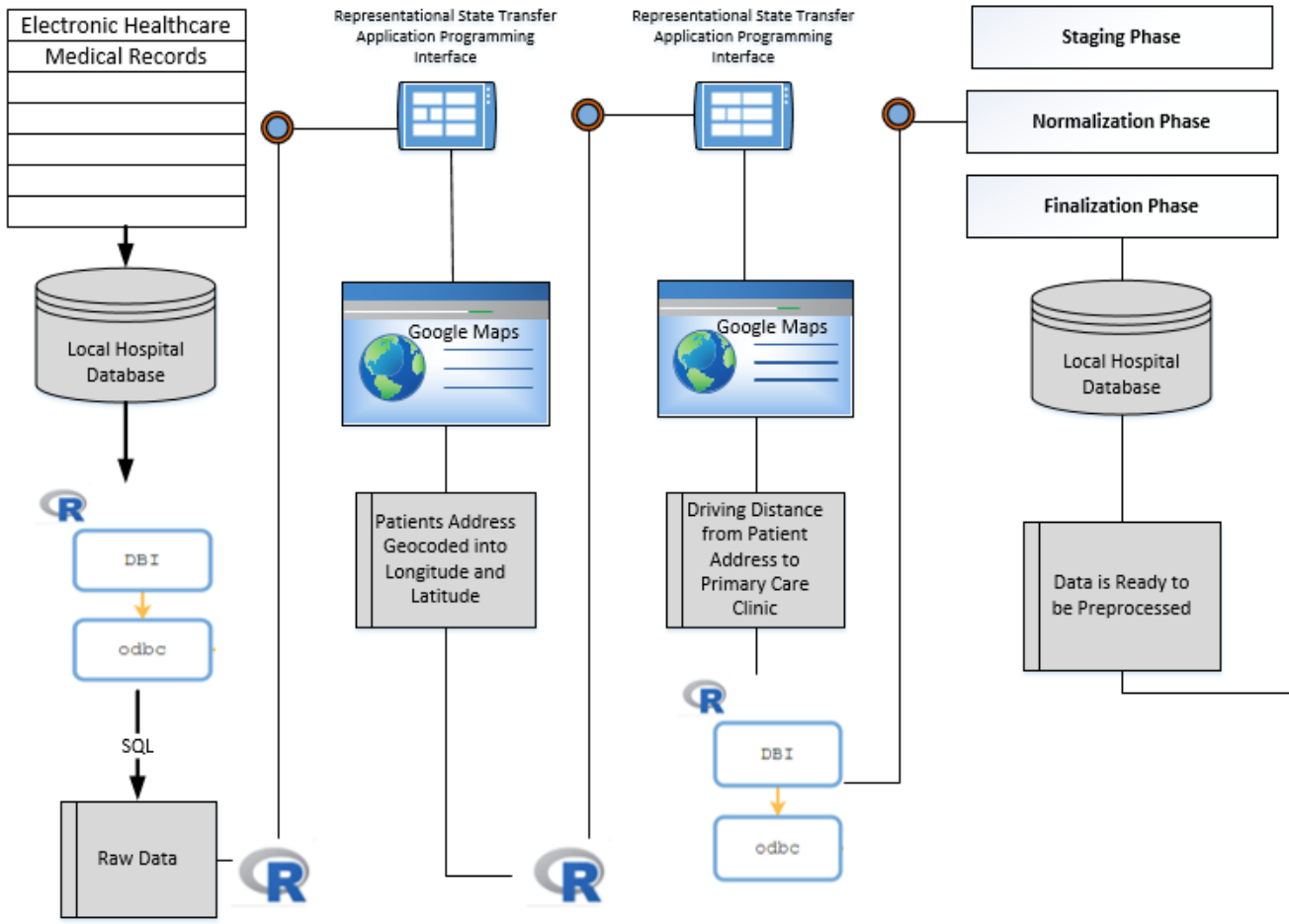




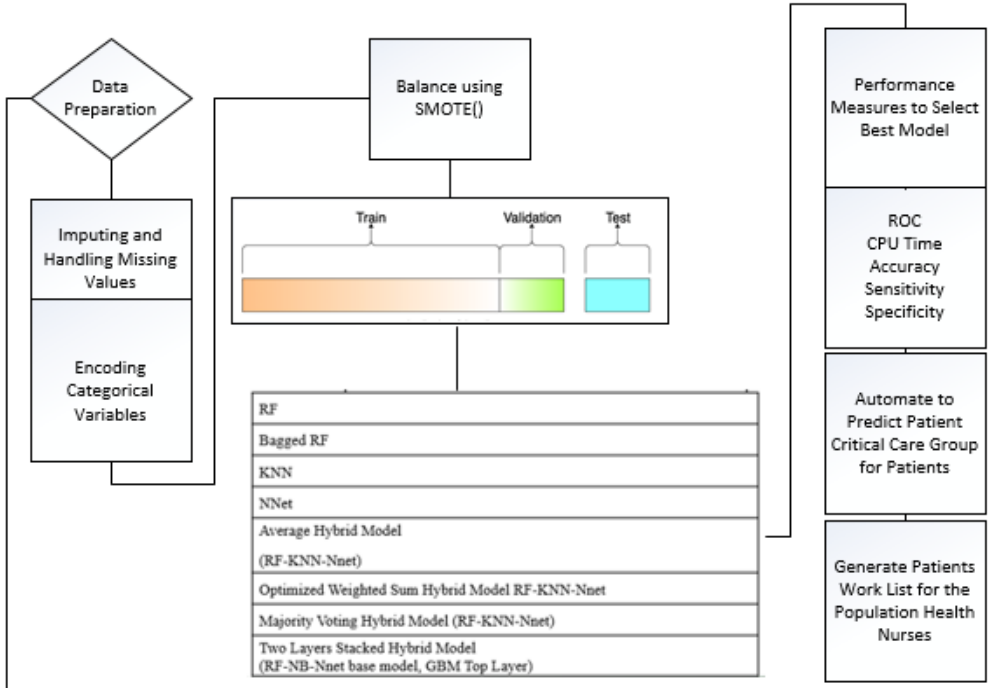
** Distance Group in Miles
A: [0,5]
B: (5,15]
C: > 15

Proposed Integrated Framework

Data Integration, Staging, Normalization, and Finalization

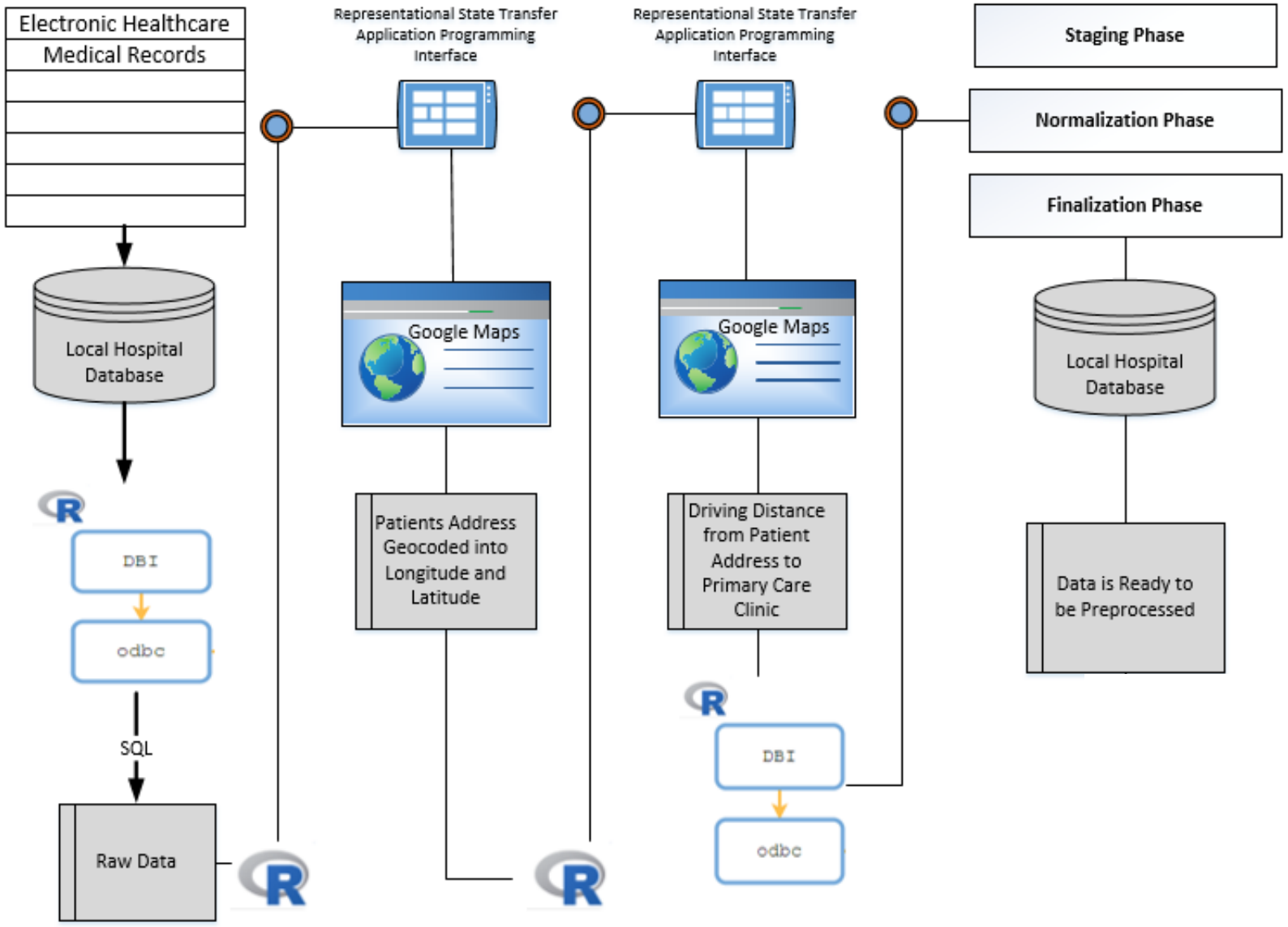


Preprocessing, Training, Testing, Automation



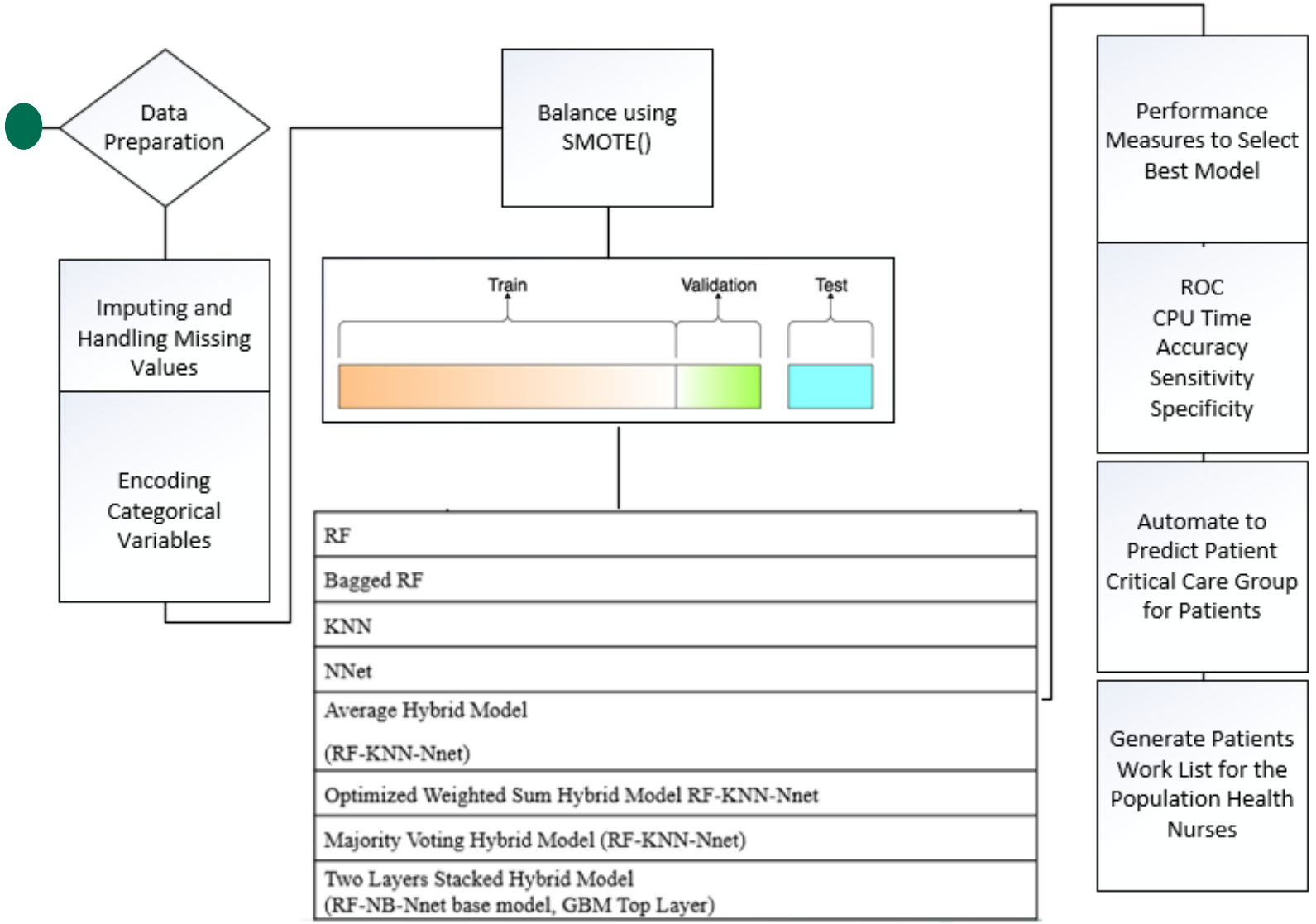
SMOTE: synthetic minority over-sampling technique

Data Acquisition, Staging, Normalization, and Finalization

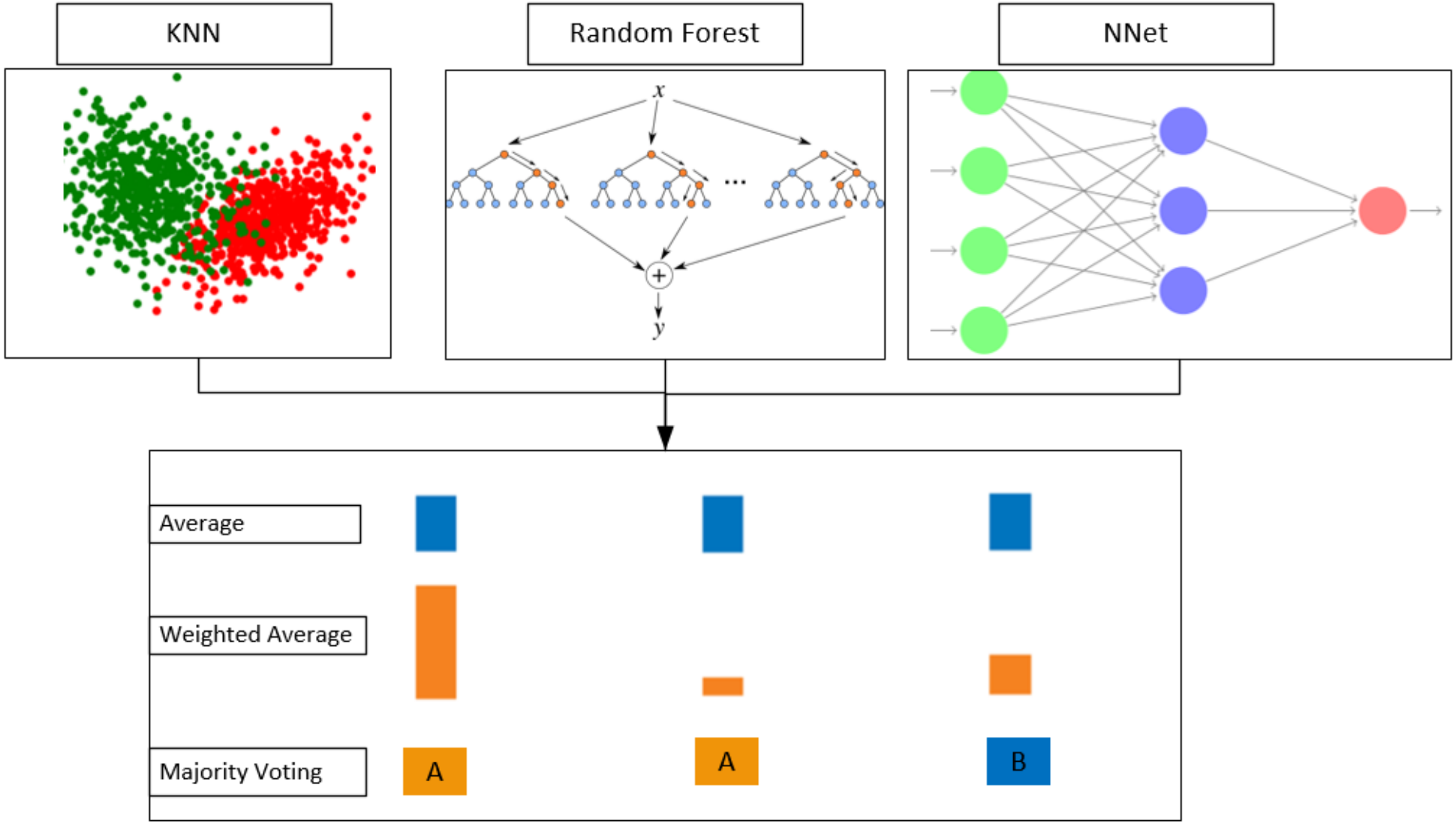


DBI : Database Interface
ODBC: Microsoft Open Database Connectivity

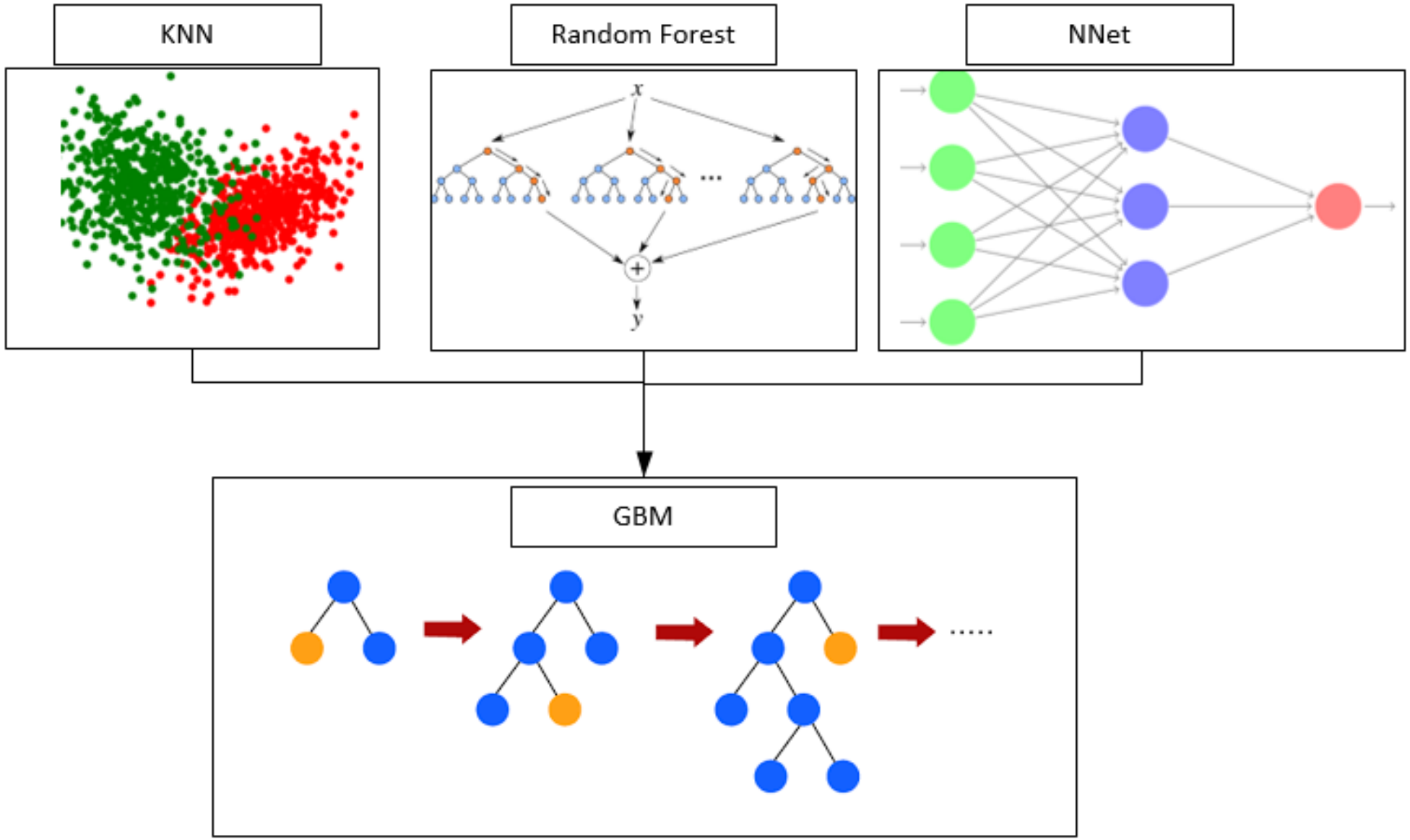
Machine Learning Framework



Hybrid Models (1/2)

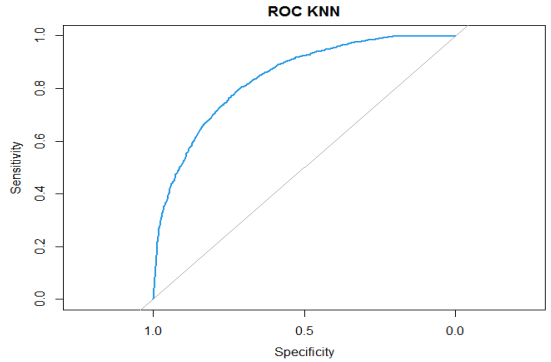


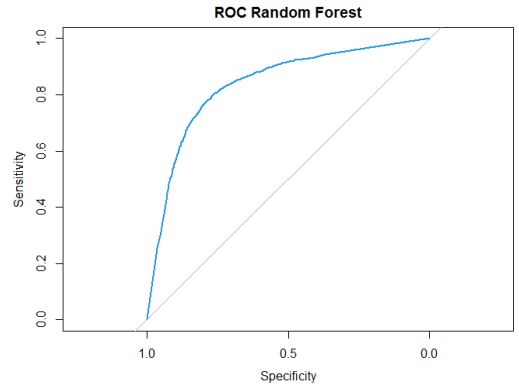
Hybrid Models (2/2)

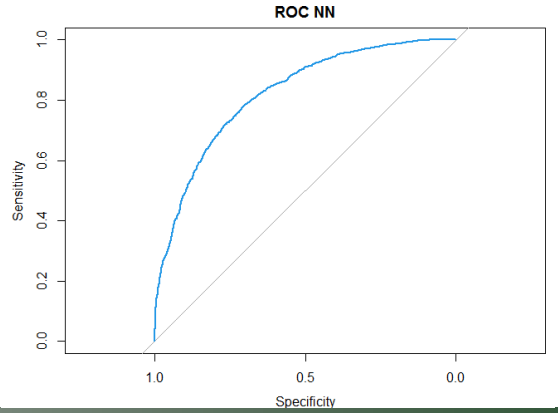


Performance and Computational Cost Metrics

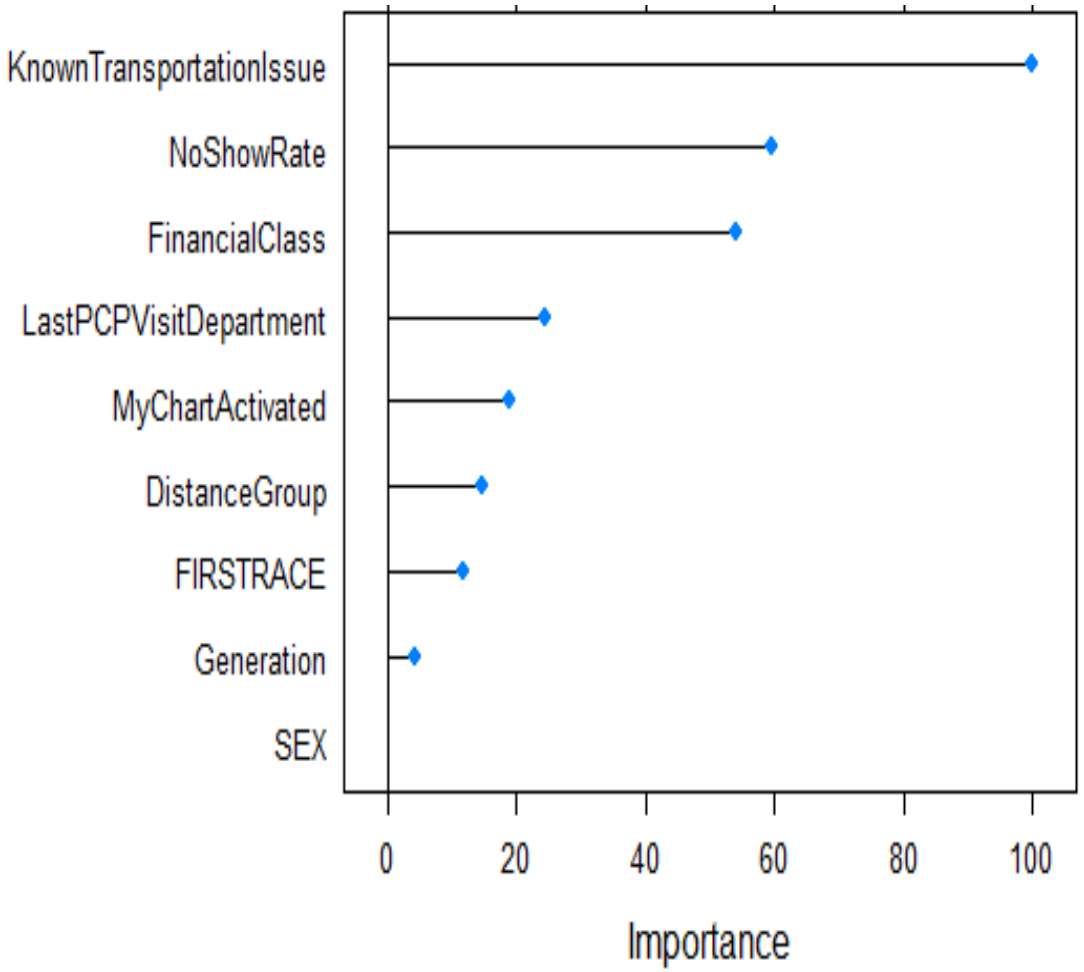
Model	Accuracy	Sensitivity	Specificity	Normalized CPU Time (min)
RF	74.28%	70.3%	78.21%	37
Bagged RF	74.47%	72.1%	77.8%	20
KNN	76%	78%	73%	16
FF-Nnet	74%	77%	71%	8
Average Hybrid Model_RF-KNN-Nnet	78%	79%	76%	> 60
Optimized Weighted Sum Hybrid Model_RF-KNN-Nnet	77%	79%	75%	> 60
Majority Voting Hybrid Model_RF-KNN-Nnet	78%	79%	77%	> 60
Two Layers Stacked Hybrid Model (RF-NB-Nnet base model, GBM Top Layer)	82%	79%	80%	> 60







Variable Importance



Factor Level	Overall
KnownTransportationIssue-Yes	100
NoShowRateHigh	91.8
DistanceGroupB	28.688
SEX Male	23.8
MyChartActivated-Yes	21.792
Clinic A	18.277
FIRSTRACEWhite	17.291
DistanceGroupC	16.662
FinancialClassMedicaidManaged	12.528
GenerationGenerationX	12.452
GenerationMillenials	12.349
Clinic B	10.495
Financial Class Medicare Managed	10.178
Clinic C	10.105

Results, Conclusion, and Application Potential

- Developed Integrated AI-driven frameworks specific to marginalized patient populations (Under researched Area)
- Developed an automated AI application that generates a list of patients' names who could benefit from early interventions to help them attend their primary wellness visit as well as reduce their critical care use.
- The study inspected for the first time the effect of distance decay and other social determinants
- Meta-ensembles and hybrid predictive models.
- Multiple models were evaluated for generalizing capability for the used dataset
- The Hybrid Model based on (RF-NB-Nnet, GBM Top Layer) model showed a potential to identify contributing factors to poor health outcomes, such as unavailability of transportation, no-show rate, insurance type, clinic, and travel distance.

Future Directions

- Improve data representation
  - Develop AI algorithms that are trained on diverse data sets and specifically designed to address the unique needs and characteristics of marginalized patient populations.
- Increase access to healthcare services through AI-enabled technologies
  - Transportation broader coverage , telemedicine and chatbots, Patient work List to propose alternatives
- Address Social determinants of health.
  - Developing AI models that can identify and address social determinants of health to improve healthcare outcomes for marginalized patients.

Publications and Conferences

- Induction of Labor from Decision to Delivery Using Hybrid Predictive Models based on a Metaheuristic Feature Selection Approach

Anemone Kasasbeh1, Liliane El-Kassis, and Hiroki Sayama (Ready to Submit)

International Journal of Clinical Practice

- Web-based Healthcare Delivery Integrated System to Forecast COVID-19 Hospitalizations in a Marginalized Patient Population: A Case Study in Broome County, New York,

A Kasasbeh, M Yildirim, A Booth, N Khan, H Sayama (Submitted)

Journal of Environmental and Public Health

- Modelling the Impact of Transportation Availability and Travel Distance on Healthcare Outcomes: A Bagged Random Forest Approach

A Kasasbeh, M Yildirim, A Booth, N Khan, H Sayama

IISE Annual Conference Proceedings (2023)

- Influential Factors for Failure to Show up for a Postpartum Visit

A Kasasbeh, M Yildirim, A Booth, MT Khasawneh

IISE Annual Conference & Expo 2019, 883-889 (2019)

- Crash severity prediction using a series of artificial neural networks

A Kasasbeh, R Shabbar, D Santos

IISE Annual Conference. Proceedings, 443-448 (2018)

- Charging station allocation for electric vehicle network using stochastic modeling and grey wolf optimization

R Shabbar, A Kasasbeh, MM Ahmed

Sustainability 13 (6), 3314 (2021)

- Crash Analysis Using Artificial Neural Network and Decision Tree

A Kasasbeh, R Shabbar

Industrial and Systems Engineering Review (ISER) (2017)

- Demand forecasting for inventory control: A case study on automotive spare parts in Saudi Arabia

N Khan, A Kasasbeh, R Alkhasawneh

- IISE Annual Conference & Expo 2018 (2018)
- Proactive Event Management using ANN with PSO Prediction in Transport Processes

THANK YOU!

Questions?