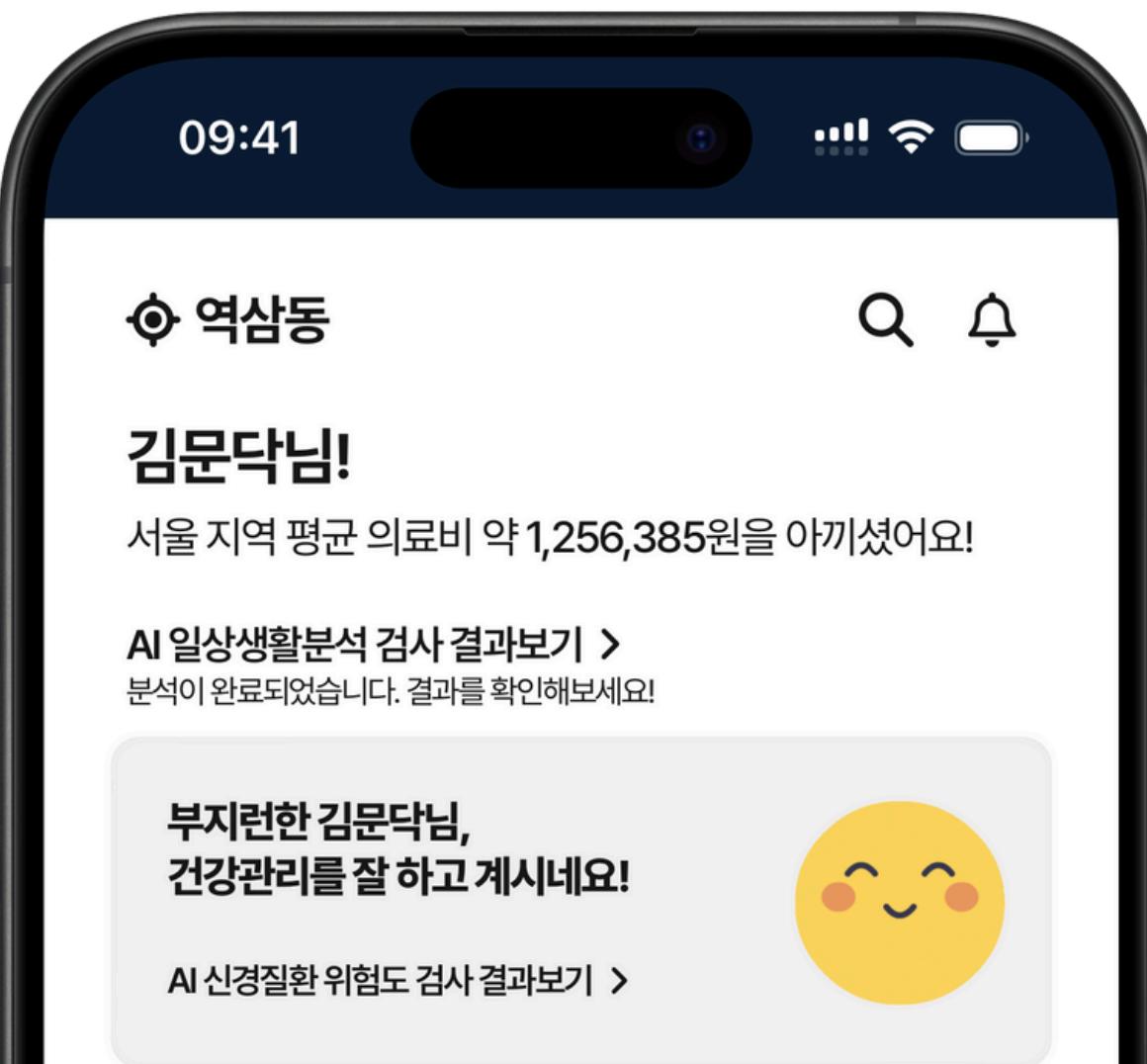




Product Development Presentation

AI Health Risk Prediction and Personalized Rehabilitation Integration Solution for
Neurological Patients



Doctors out the Door

정혜원 박지건 윤지훈 이상훈 임예원 황유진

11-06-2024

Index

01 **Background of Planning**

Topic Selection / Core Features / Service Goals / Open Data

02 **Data Analysis**

Data Collection / Variable Selection / ERD / Data Wrangling / EDA

03 **Machine Learning**

Preprocessing / Modeling / Evaluation / Interpretation

04 **Multi Platform Development**

Overview / Technologies / Goals and Tasks / Operational plan

04 **Business Model**

Market Analysis / Target Analysis / Expected Outcomes

Competition

Team Members



Startup Competition using Healthcare Big Data

Hosted by: Ministry of Health and Welfare

Organized by: Health Insurance Review and Assessment Service, National Health Insurance Service

Competition Category: Product and Service Development

Project Period : 25.4.2024 ~ 31.5.2024

Competition

Team Members



Hyewon Jung
PM



Jigeon Park
Full-stack Developer



Jihoon Youn
Service Planner



Sanghoon Lee
Front-end Developer



Yewon Lim
Data Analyst

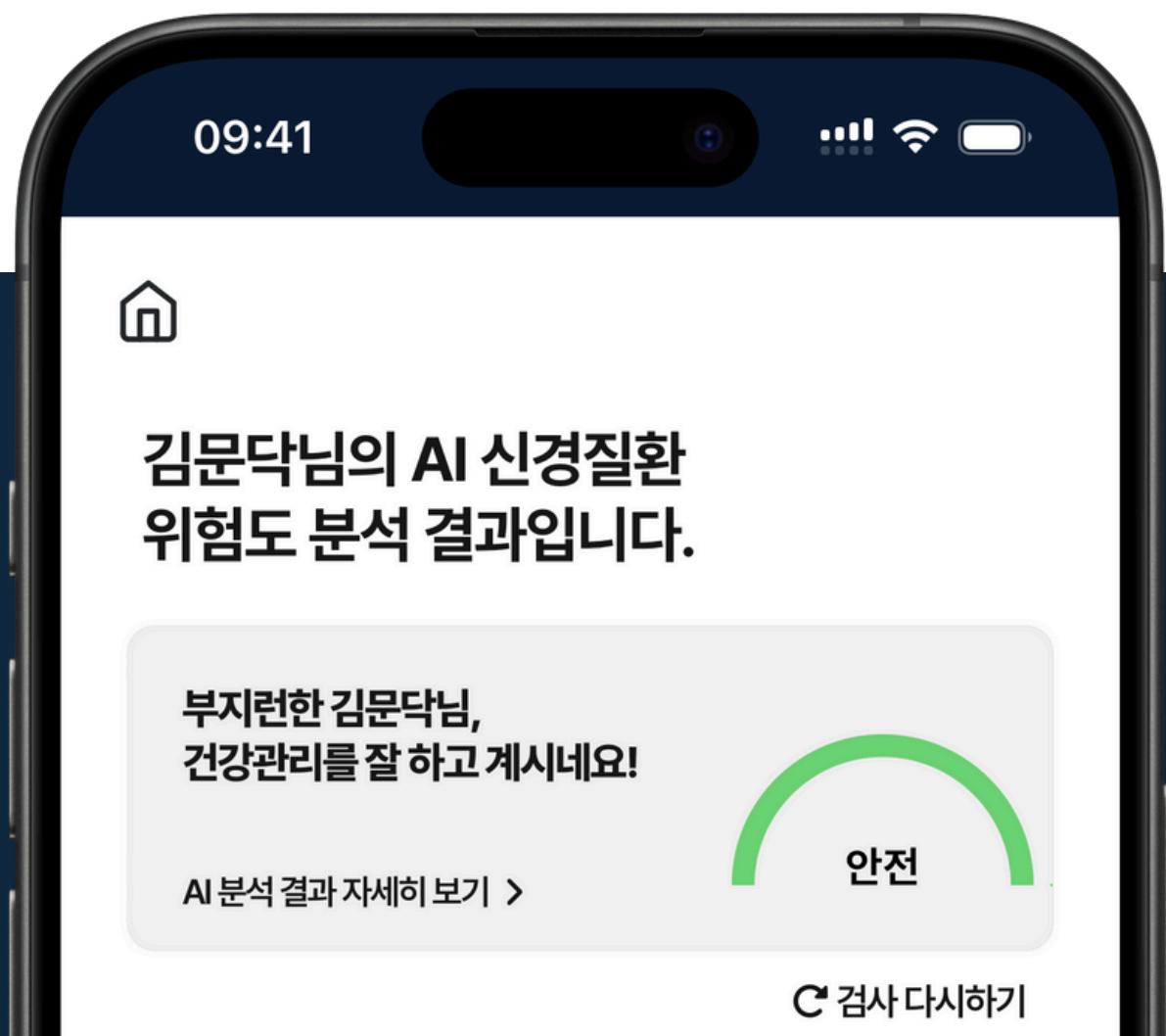


Youjin Hwang
Data Scientist



Background of Planning

Topic Selection / Core Features / Service Goals / Open Data



Planning Dept
Hyewon Jung (PM)
Jihoon Youn (Planning))

Topic Selection

Core Features

Service Goals

Open Data

Business Problem

N 뇌출혈 재활

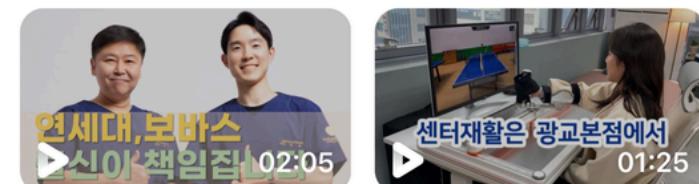
블로그 카페 이미지 지식iN

파워링크 뇌출혈재활 관련광고 ⓘ

더건강연세재활센터 · 재활센터.com 광고
(주)더건강연세재활운동센터 · 가장 소중한 내 가족의 재활 · 진심을 다해 재활 하겠습니다

뇌출혈재활 전국18개 지점, 서울에서 제주까지 모든 지역 방문재활 가능 세상에서 가장 소중한 내 가족을 위해 정성을 다해 방문치료 & 재활치료 하겠습니다

첫 방문 무료평가 >



더건강연세 홍보영상

myongji-sm.co.kr/ 광고
뇌혈관전문병원, 명지성모병원

N 뇌경색 재활

블로그 카페 이미지 지식iN

파워링크 뇌경색재활 관련광고 ⓘ

mjchoonhey.co.kr/ 광고 톡톡
재활전문병원 명지춘혜병원
보건복지부지정 재활전문병원, 첨단 로봇재활치료, 운동치료, 작업치료, 언어치료

진료시간안내 진료과 진료예약

heavenlyhospital.com 광고

사랑을 실천하는 해븐리병원
신경과 전문의 6인, 두통, 어지럼증, 이명, 불면증, 수면다원검사
의138261

lifeclinic.co.kr 광고

라이프의원 양재역 4번출구
서울대종양 내과저문의 기선구 신자 겸도매 간사서

N 뇌졸중 재활

블로그 카페 이미지 지식iN

파워링크 뇌졸중재활 관련광고 ⓘ

purunhospital.com 광고
청주 재활, 씨엔씨푸른병원
회복기재활, 뇌졸중재활중심, 간호간병통합서비스, 건강증진센터
의166705

the-jain.co.kr 광고

재활치료는 더자인병원
근골격계 질환, 척추손상, 뇌졸중, 파킨스병, 외상성 뇌손상, 중추신경 손상
의157353

mjchoonhey.co.kr/ 광고 톡톡

재활전문병원 명지춘혜병원
보건복지부지정 재활전문병원, 첨단 로봇재활치료, 운동하고, 잠이하고, 일어나하고

N 파킨슨 재활

블로그 카페 이미지 지식iN

파워링크 파킨슨재활 관련광고 ⓘ

hroboticsinc.com/ 광고 pay+
파킨슨재활 에이치로보틱스 · 홈 재활, 7일 무료 체험!

집에서 쉽게 뇌졸중 편마비, 근골격계 재활치료가 가능한 리블레스



렌탈신청 7일무료체험 리퍼브상품

올바른재활.com 광고

파킨슨재활은 올바른재활 · 365일 실시간 상담 가능

임상경력 10년이상 재활전문가들, 1:1프리미엄 맞춤 재활, 파킨슨재활, 파킨슨재활



Topic Selection

Core Features

Service Goals

Open Data

Business Problem

Too many Hospital ads,

N 뇌출혈 재활

블로그 카페 이미지 지식iN

파워링크 뇌출혈재활 관련광고 ②

더건강연세재활센터 · 재활센터.com 광고
(주)더건강연세재활운동센터 · 가장 소중한 내 가족의 재활 · 진심을 다해 재활 하겠습니다

뇌출혈재활 전국18개 지점, 서울에서 제주까지 모든 지역 방문재활 가능 세상에서 가장 소중한 내 가족을 위해 정성을 다해 방문치료 & 재활치료 하겠습니다

첫 방문 무료평가 >
▶ 세대교체마스▶신아책임집 02:05
▶센터재활은 광교분점에서 01:25

더건강연세 홍보영상

myongji-sm.co.kr/ 광고
뇌혈관전문병원, 명지성모병원

N 뇌경색재활

블로그 카페 이미지 지식iN

파워링크 뇌경색재활 관련광고 ②

mjchoonhey.co.kr/ 광고 □ 특록
재활전문병원 명지춘혜병원
보건복지부지정 재활전문병원, 첨단 로봇재활치료, 운동치료, 작업치료, 언어치료

진료시간안내 진료과 진료예약

heavenlyhospital.com 광고
사랑을 실천하는 해븐리병원
신경과 전문의 6인, 두통, 어지럼증, 이명, 불면증, 수면다원검사

의138261

lifeclinic.co.kr 광고
라이프의원 양재역 4번출구
서울대종합내과저문의 기선구 심자 겸도매 간사서

N 뇌졸중재활

블로그 카페 이미지 지식iN

파워링크 뇌졸중재활 관련광고 ②

purunhospital.com 광고
청주 재활, 씨엔씨푸른병원
회복기재활, 뇌졸중재활중심, 간호간병통합서비스, 건강증진센터
의166705

the-jain.co.kr 광고
재활치료는 더자인병원
근골격계 질환, 척추손상, 뇌졸중, 파킨스병, 외상성 뇌손상, 중추신경 손상
의157353

mjchoonhey.co.kr/ 광고 □ 특록
재활전문병원 명지춘혜병원
보건복지부지정 재활전문병원, 첨단 로봇재활치료, 운동치료, 작업치료, 언어치료

N 파킨슨 재활

블로그 카페 이미지 지식iN

파워링크 파킨슨재활 관련광고 ②

hroboticsinc.com/ 광고 □ pay+
파킨슨재활 에이치로보틱스 · 홈 재활, 7일 무료 체험!

집에서 쉽게 뇌졸중 편마비, 근골격계 재활치료가 가능한 리블레스



렌탈신청 7일무료체험 리퍼브상품
▶ 을바른재활.com 광고

파킨슨재활은 을바른재활 · 365일 실시간 상담 가능

임상경력 10년이상 재활전문가들, 1:1프리미엄 맞춤 재활, 파킨슨재활, 파킨슨재활



Business Problem

Too many Hospital ads,

Hard to find reliable information,



파워링크

(주)더건강연세재활센터 · 가장 소중한 내 가족의 재활 · 진심을 다해 재활 하겠습니다
뇌출혈재활 전국18개 지점, 서울에서 제주까지 모든 지역 방문재활 가능 세상에서 가장 소중한 내 가족을 위해 정성을 다해 방문치료 & 재활치료 하겠습니다

▶ 첫 방문 무료평가 >

▶ 세대교체마스▶신이책임집 02:05
▶센터재활은 광교분점에서 01:25

더건강연세 홍보영상

뇌출혈재활센터 · 뇌출혈재활센터.com 광고
mjchoonhey.co.kr/ 광고 □ 특록

재활전문병원 명지춘혜병원
보건복지부지정 재활전문병원, 첨단 로봇재활치료, 운동치료, 작업치료, 언어치료

진료시간안내 ▶ 진료과 ▶ 진료예약

▶ heavenlyhospital.com 광고
사랑을 실천하는 해븐리병원
신경과 전문의 6인, 두통, 어지럼증, 이명, 불면증, 수면다원검사
의138261

▶ myongji-sm.co.kr/ 광고
뇌혈관전문병원, 명지성모병원
서울대종양내과저문의 기선구 심자 겸도매 간사서

파킨슨재활 관리센터 · mjchoonhey.co.kr/ 광고 □ 특록
청주 재활, 씨엔씨푸른병원
회복기재활, 뇌졸중재활중심, 간호간병통합서비스, 건강증진센터
의166705

▶ the-jain.co.kr 광고
재활치료는 더자인병원
근골격계 질환, 척추손상, 뇌졸중, 파킨스병, 외상성 뇌손상, 중추신경 손상
의157353

▶ mjchoonhey.co.kr/ 광고 □ 특록
재활전문병원 명지춘혜병원
보건복지부지정 재활전문병원, 첨단 로봇재활치료, 운동치료, 작업치료, 언어치료

파킨슨재활 관리센터 · hroboticsinc.com/ 광고 □ 특록
파킨슨재활 에이치로보틱스 · 홈 재활, 7일 무료 체험!
집에서 쉽게 뇌졸중 편마비, 근골격계 재활치료가 가능한 리블레스



▶ 을바른재활.com 광고
렌탈신청 ▶ 7일무료체험 ▶ 리퍼브상품

▶ 을바른재활.com 광고
파킨슨재활은 을바른재활 · 365일 실시간 상담 가능
임상경력 10년이상 재활전문가들, 1:1프리미엄 맞춤 재활, 파킨슨재활, 파킨슨재활



Business Problem

Too many Hospital ads,

Hard to find reliable information,

Information available is scattered



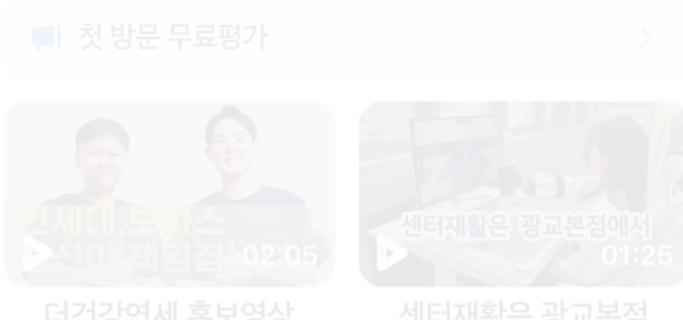
파워링크

(주)더건강연세재활운동센터 · 가장 소중한 내 가족의 재활 · 진심을 다해 재활 하겠습니다
뇌출혈재활 전국18개 지점, 서울에서 지역 방문재활 가능 세상에서 가장 소중한 내 가족을 위해 정성을 다해 방문치료 & 재활치료 하겠습니다

뇌출혈재활센터 · 가장 소중한 내 가족의 재활
보건복지부지정 재활전문병원, 첨단 로봇재활치료,
진료시간연대
진료과
진료예약

재활전문병원 명지춘혜병원
청주 재활, 씨엔씨푸른병원
회복기재활, 뇌졸중재활중심, 간호간병통합서비스,
재활치료는 더자인병원

파킨슨재활 에이치로보틱스 · 흄 재활, 7일 무료 체험!
파킨슨재활은 올바른재활 · 365일 실시간 상담 가능
임상경력 10년이상 재활전문가들, 1:1프리미엄 맞춤 재활, 파킨슨재활, 파킨슨재활



첫 방문 무료평가
▶ 세대교체
▶ 신아책임집 02:05
▶ 센터재활은 광교분점에서 01:25
더건강연세 홍보영상
센터재활은 광교분점

the-jain.co.kr 광고
재활치료는 더자인병원
근골격계 질환, 척추손상, 뇌졸증, 파킨스병, 외상성 뇌손상, 중추신경 손상
의157353

렌탈신청
7일무료체험
리퍼브상품
▶ 올바른재활.com 광고



myongji-sm.co.kr 광고
뇌혈관전문병원, 명지성모병원
라이프의원 양재역 4번출구
서울대종합내과저문의 기선구 심자 겸도매 가사서

mjchoonhey.co.kr/ 광고
재활전문병원 명지춘혜병원
보건복지부지정 재활전문병원, 첨단 로봇재활치료,
온드하고, 잠업하고, 일어나고



Business Problem

Too many Hospital ads,

Hard to find reliable information,

Information available is scattered,

Rehabilitation is urgent

Business Problem



What's my **current** status?

Business Problem



What's my **golden time**?

Business Problem



What kind of
rehabilitation do I need?

Legislative Amendment

Policy > institutions and laws

'Visit rehabilitation treatment' will start next year...visiting rehabilitation fee up to 180,000 won



Reporter Lee Ji-hyun.

Issue date: 2022-11-23 18:40:08



Implementation of a pilot project in the third phase of rehabilitation medical institutions...Severe among discharged patients Service for 90 days after discharge...Allow extension of 30 days for patient condition consideration Incentives to activate links to medical institutions in the acute phase of cerebrovascular disease → palliative phase

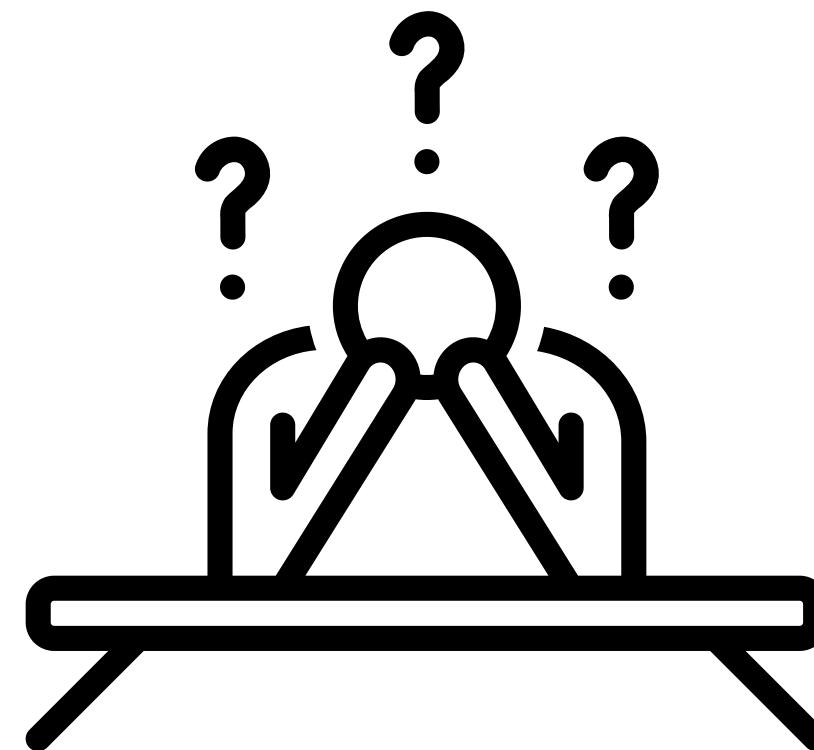
[Medical Times = Reporter Lee Ji-hyun] The Ministry of Health and Welfare will promote a two-year pilot project for visiting rehabilitation treatment from January next year (2023).

If the number of rehabilitation medical institutions has been targeted at inpatients as the next stage of the second phase of the pilot project, the third phase will be targeted at patients who need home rehabilitation among discharged patients.

On the 23rd, the Ministry of Health and Welfare held the 22nd Health Insurance Policy Review Committee (hereinafter referred to as the Health Insurance Review Committee) and reported on the third phase of the pilot project (visiting rehabilitation treatment pilot project).

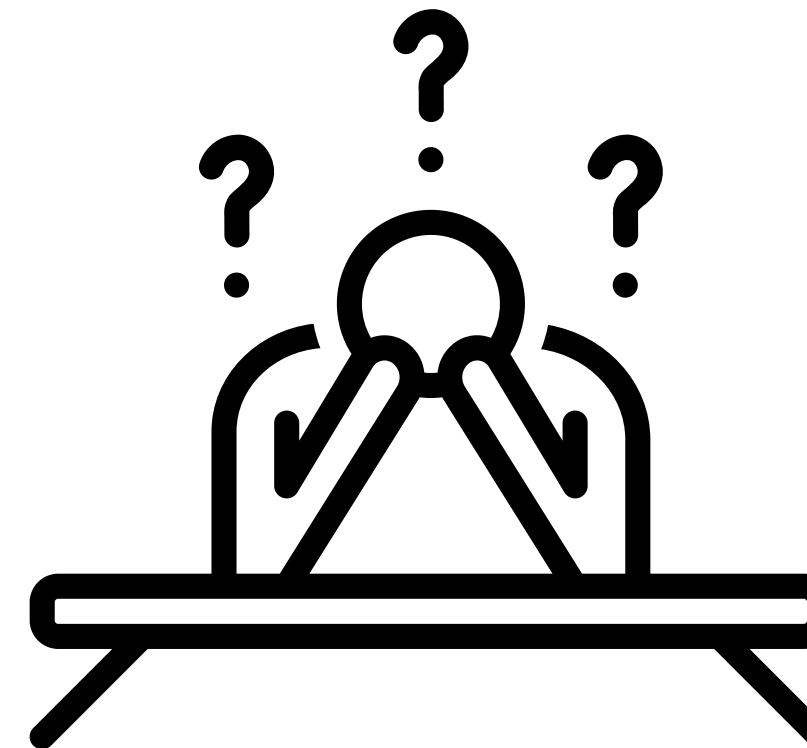
■ How is the door-to-door rehabilitation treatment service conducted?

Legislative Amendment



For patients with **mobility difficulties**

Legislative Amendment



Providing the **first application service** that connects
home rehabilitation for patients with mobility difficulties.

Topic Selection

**AI Health Risk Prediction and Personalized Rehabilitation Integration
Solution for Neurological Patients**

Core Features

AI Risk Prediction

**Personalised
Rehabilitation
Integration**

**1:1
Home
Rehabilitation
Matching**

Topic Selection

Core Features

Service Goals

Open Data

Service Goals



Topic Selection

Core Features

Service Goals

Open Data

Service Goals



Service Goals



Service Goals



Utilizing Open Data

HIRA Data 8

MIMIC Data 1

Kakao Real-time API 1

Utilizing Open Data

HIRA Data 8

MIMIC Data 1

Kakao Real-time API 1

A total of 10

Utilizing Open Data



건강보험심사평가원
HEALTH INSURANCE REVIEW & ASSESSMENT SERVICE

No	HIRA Open Data
1	Hospital Information Service (2024)
2	Hospital Medical Information Inquiry Service (2024)
3	Hospital Evaluation Information Service (2024)
4	Healthcare Institution Opening and Closing Information Inquiry Service (2024)
5	Outstanding Institution Hospital Evaluation Information Service (2024)
6	Detailed Information Service by Medical Institution (2024)
7	Special Treatment Hospital Information Service (2024)
8	Medical Expenses Statistics by Medical Department and Region for Hospitals and Higher (2021)

Utilizing Open Data

Database Credentialed Access

MIMIC-IV

Alistair Johnson  , Lucas Bulgarelli  , Tom Pollard  , Steven Horng  , Leo Anthony Celi  , Roger

Published: Jan. 6, 2023. Version: 2.2

Guidelines for creating datasets and models from MIMIC (April 24, 2024, 10:12 a.m.)

We recognize that there is value in creating datasets or models that are either derived from MIMIC or which augment MIMIC in some way (for example, by adding annotations). Here are some guidelines on creating these datasets and models:

- Any derived datasets or models should be treated as containing sensitive information. If you wish to share these resources, they should be shared on PhysioNet under the same agreement as the source data.
- If you would like to use the MIMIC acronym in your project name, please include the letters "Ext" (for example, MIMIC-IV-Ext-YOUR-DATASET"). Ext may either indicate "extracted" (e.g. a derived subset) or "extended" (e.g. annotations), depending on your use case.

When using this resource, please cite: [\(show more options\)](#)

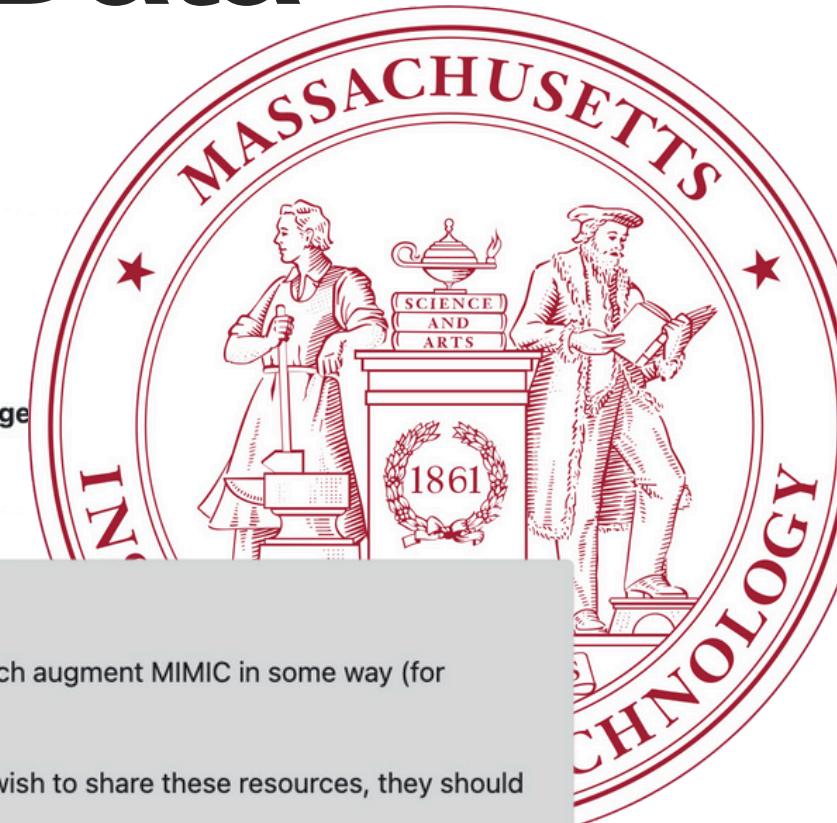
Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L. A., & Mark, R. (2023). MIMIC-IV (version 2.2). PhysioNet. <https://doi.org/10.13026/6mm1-ek67>.

Additionally, please cite the original publication:

Johnson, A.E.W., Bulgarelli, L., Shen, L. et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data* 10, 1 (2023). <https://doi.org/10.1038/s41597-022-01899-x>

Please include the standard citation for PhysioNet: [\(show more options\)](#)

Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* [Online]. 101 (23), pp. e215–e220.



Contents ▾

Share



Access

Access Policy:

Only credentialed users who sign the DUA can access the files.

Topic Selection

Core Features

Service Goals

Open Data

Utilizing Open Data

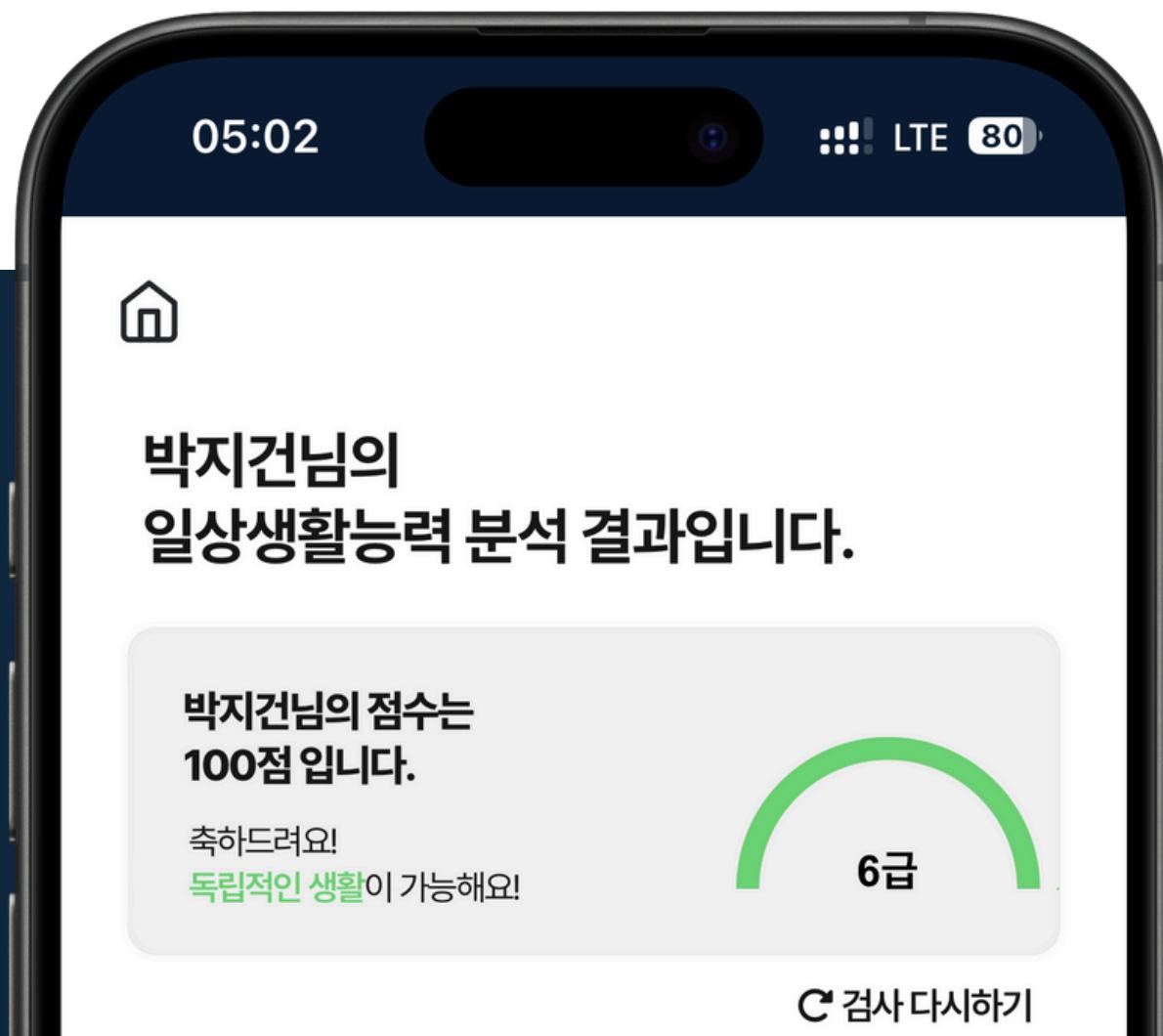
Kakao Postcode API





Data Analysis

Data Collection / Variable Selection / ERD / Data Wrangling / EDA



Data Analysis Dept.
Yewon Lim (Data Analyst)

Data Collection



A large database for over
40,000 patients

Jointly developed by
MIT Research Group &
BIDMC

Applied in **various fields**
such as medical data
science and AI

Data Collection



Completion Date 01-May-2024
Expiration Date 01-May-2027
Record ID 62532200

This is to certify that:

JIGEON PARK

Has completed the following CITI Program course:

Human Research
(Curriculum Group)
IRB Members
(Course Learner Group)
2 - Refresher Course
(Stage)

Under requirements set by:

Massachusetts Institute of Technology Affiliates

CITI
Collaborative Institutional Training Initiative

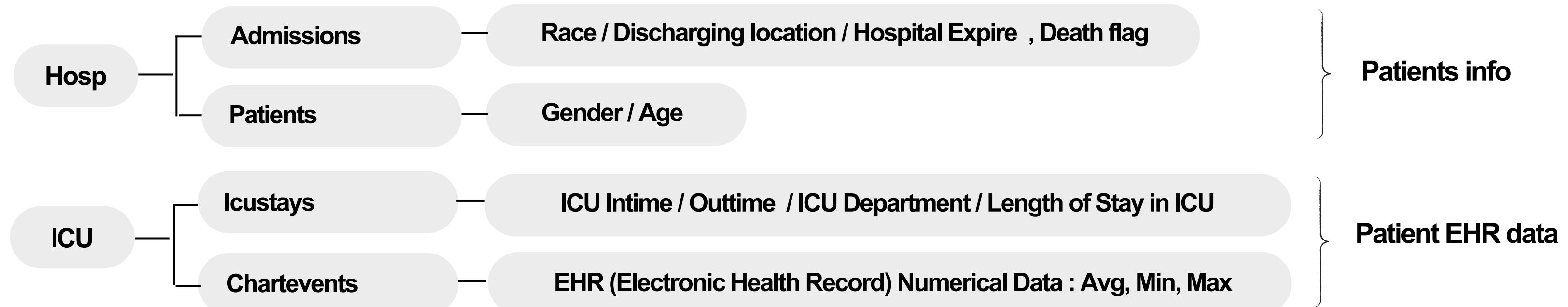
101 NE 3rd Avenue, Suite 320
Fort Lauderdale, FL 33301 US
www.citiprogram.org

Generated on 01-May-2024. Verify at www.citiprogram.org/verify/?w79f8cde5-0b32-4b37-afa0-836e0ac96f0c-62532200

Connect to Google
BigQuery

Variable Selection

Tables / Columns for Analysis



Variable Selection

Predictor variables / Target

X

Y

	Chartevents		Race	Age	Gender	LOS	Death
	No.	Numerical : 4	Categorical : 7				
Patient 1							
Patient 2			
.			
Patient 1731							Died : 1 Alive : 0

Variable Selection

EHR Items

Chartevents

- GCS (Verbal Response, Eye Opening, Motor Response)
- Richmond RAS Scale
- Gait Transferring
- Mental Status
- Secondary Diagnosis
- Heart Rate
- Respiratory Rate
- Blood Pressure (Diastolic, Systolic)



Considering the counts of each item, user convenience, and the number of null values, 11 variables have been selected

Variable Selection

EHR Ordinal Variables

Variable	Details
GCS Verbal Response	<ul style="list-style-type: none">• Glasgow Coma Scale (GCS) - Verbal Response Item• Purpose: Assesses the patient's level of consciousness.• Scoring: Ranges from 1 point (no response) to 5 points (normal response).
GCS Eye Opening	<ul style="list-style-type: none">• Glasgow Coma Scale - Eye Opening Response Item• Scoring: Ranges from 1 point (no response) to 4 points (spontaneous eye opening)
GCS Motor Response	<ul style="list-style-type: none">• Glasgow Coma Scale - Motor Response Item• Scoring: Ranges from 1 point (no response) to 6 points (obeys commands).
Richmond RAS Scale	<ul style="list-style-type: none">• Assessment of Sedation and Agitation State• Purpose: Evaluates the patient's sedation and agitation levels.• Scoring: Ranges from -5 points (deep sedation) to +4 points (extreme agitation).

Variable Selection

EHR Numerical Variables

Variable	Details
Heart Rate	<ul style="list-style-type: none">Assesses heart condition by measuring beats per minute (bpm).Normal Range for Adults: 60-100 bpm
Respiratory Rate	<ul style="list-style-type: none">Assesses respiratory condition by counting breaths per minute.Normal Range for Adults: 12-20 breaths per minute
Blood Pressure Diastolic	<ul style="list-style-type: none">Measurement: Measures the pressure in the blood vessels when the heart is at rest.Normal Range: 60-80 mmHg
Blood Pressure Systolic	<ul style="list-style-type: none">Measures the pressure in the blood vessels when the heart contracts.Normal Range: 90-120 mmHg

Variable Selection

EHR Categorical Variables

Variable	Details
Mental Status	<ul style="list-style-type: none">• Purpose: Assesses the patient's cognitive state, including confusion, dementia, and level of consciousness.• Evaluation Criteria: Includes normal, confused, disoriented, and unresponsive states.• Scoring: Oriented to own ability (0.0) / Forgets limitations (15.0)
Secondary Diagnosis	<ul style="list-style-type: none">• Purpose: Identifies other medical conditions or diseases the patient has in addition to the primary diagnosis.• Scoring: No (0.0) / Yes (15.0)

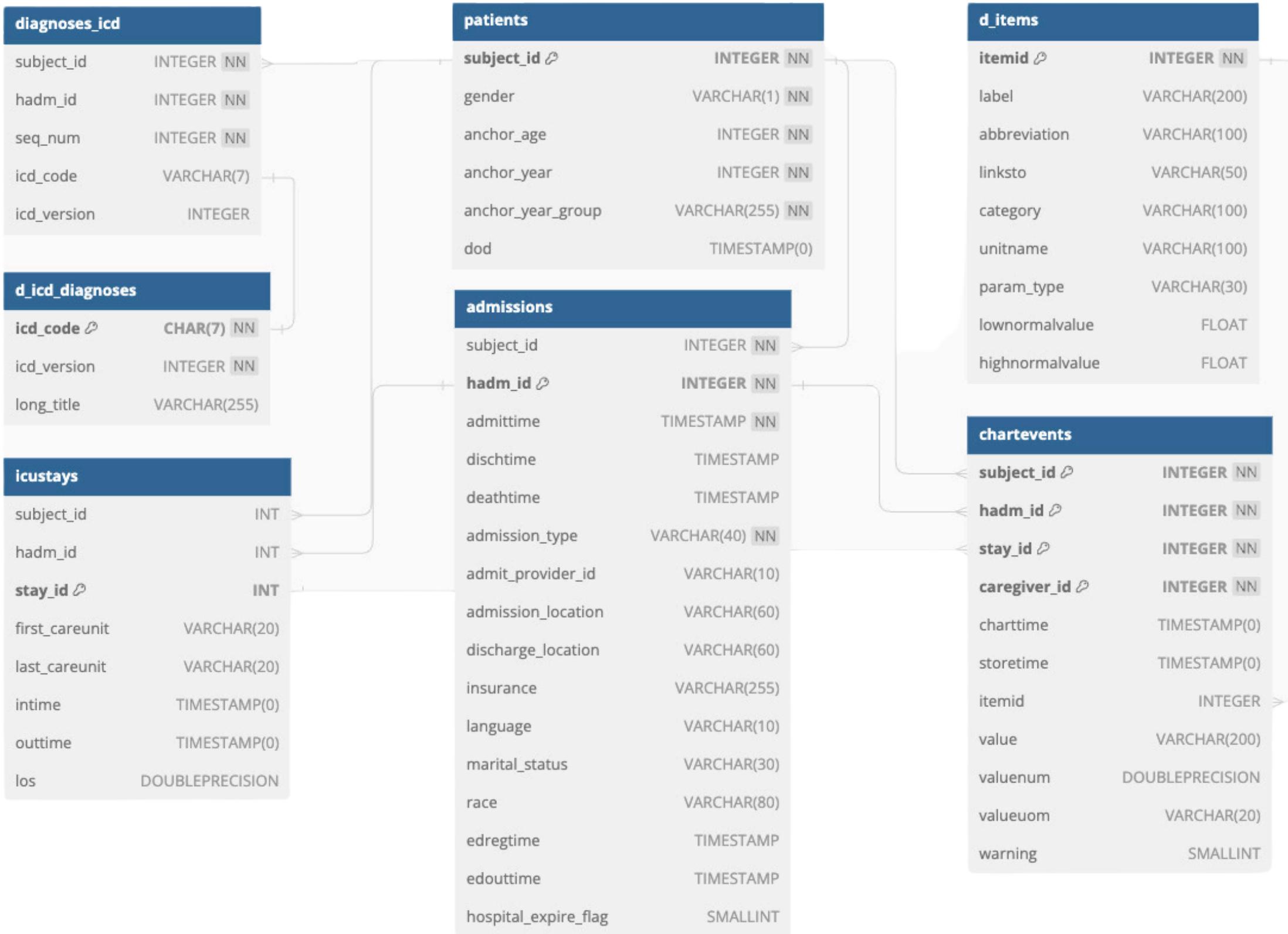
Data Collection

Variable Selection

ERD

Data Wrangling

EDA



**Creating Train Dataset by Joining
7 tables**

Data Wrangling

Data Transformation

Filtering Neuro SICU

Pivoting EHR items
from Chartevents

Creating **Target** column
(Death Flag)

Data Wrangling

Data Cleaning

Removal of Null Values

Removal of Duplicates

**Removing column with
Inaccurate Encoding in the
Original Data
(Gate/Transferring)**

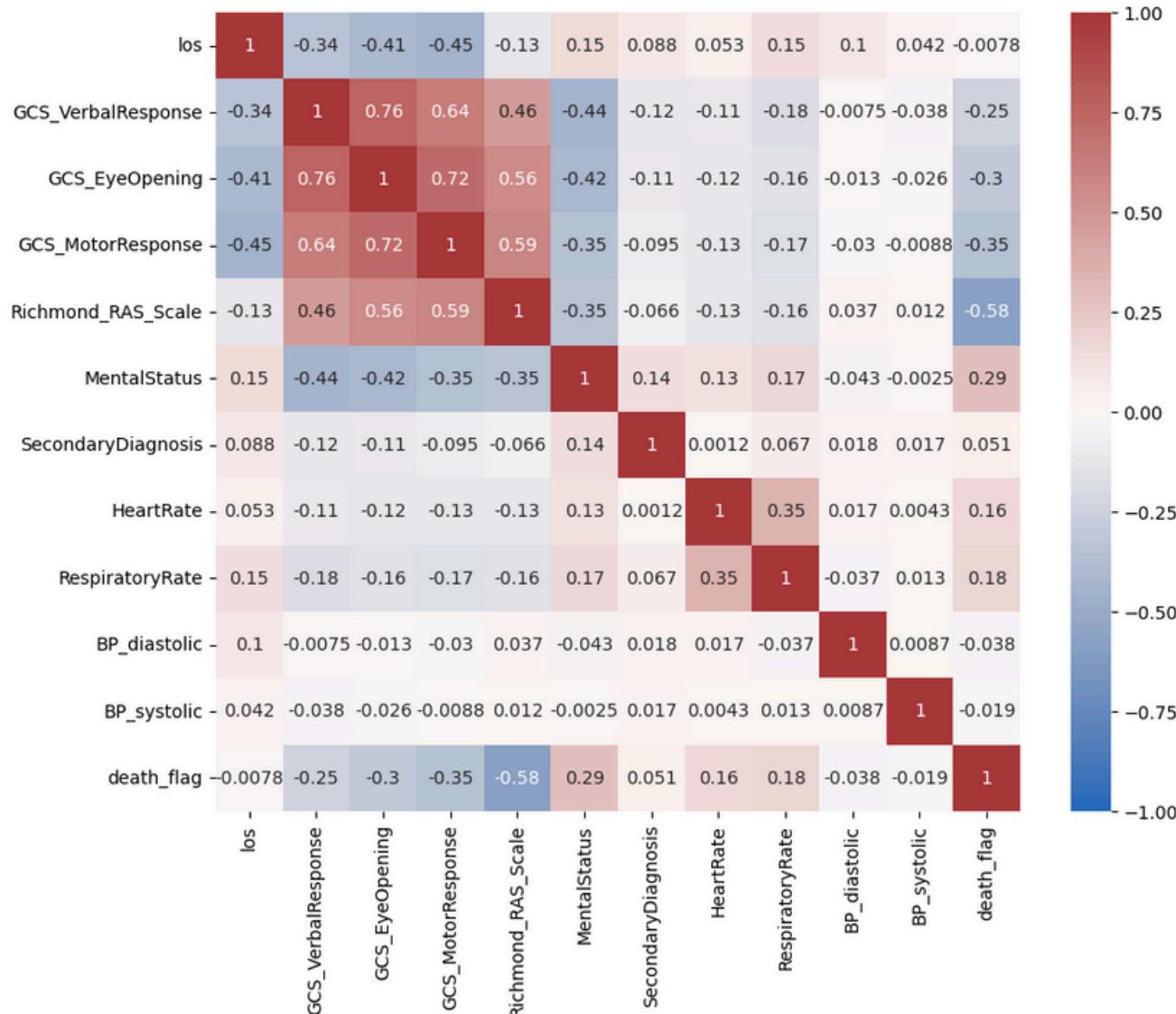
EDA

Statistical Analysis

Type	Method	Hypothesis	Result	Interpretation
Categorical (2)	Chi-Square Test	The two variables are not independent	gender, race p-value < 0.05	<ul style="list-style-type: none"> Null Hypothesis Rejected The two variables are not independent There is a correlation with the dependent variable
Numerical (11)	Independent T-Test	There is a difference in the population means of two groups	all columns except LOS p-value < 0.05	<ul style="list-style-type: none"> Null Hypothesis Rejected There is a significant difference between the means of the two groups

EDA

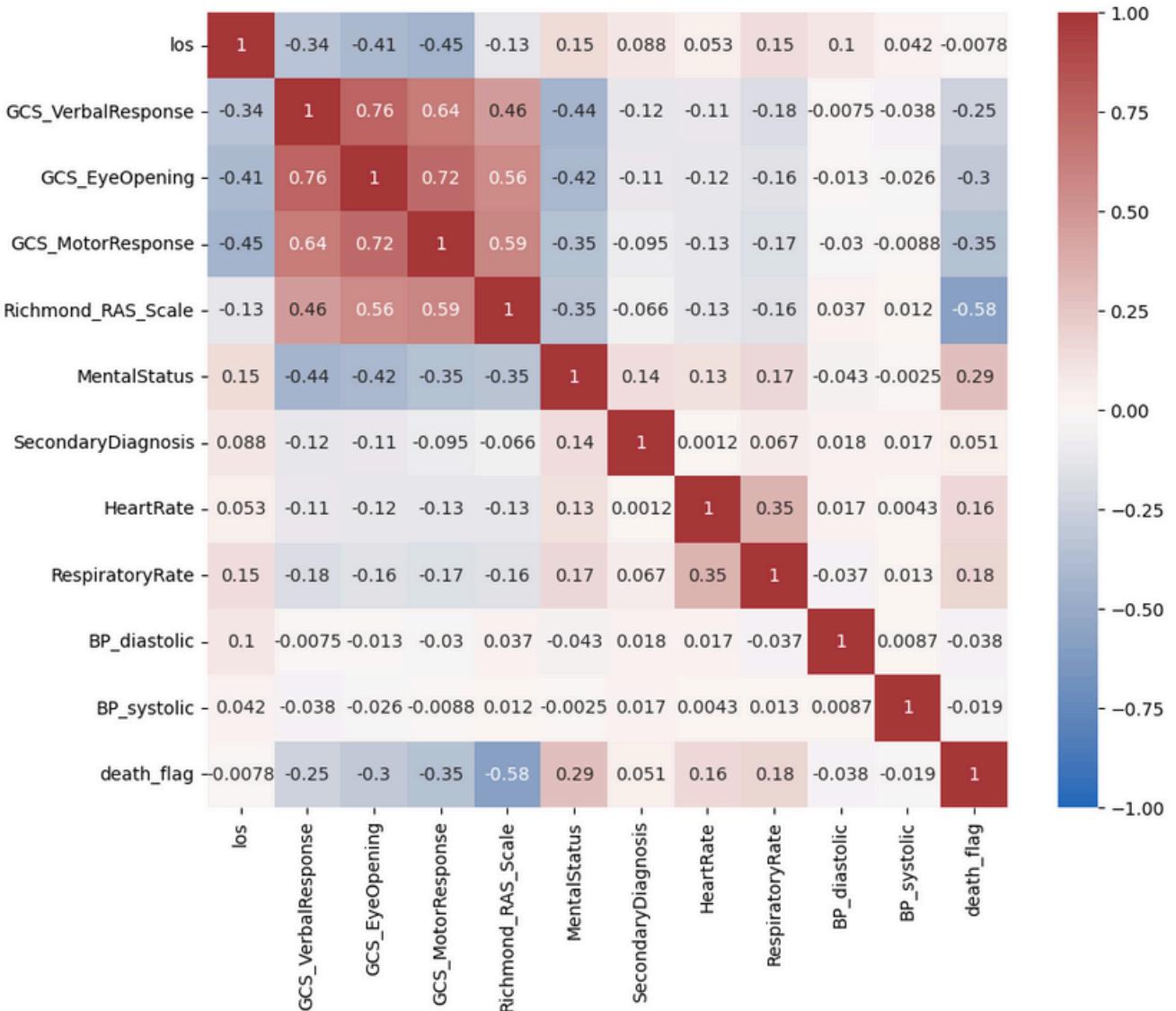
Correlation Analysis



Variable	Interpretation
GCS Related Variables	<ul style="list-style-type: none"> Negative correlation with the dependent variable. Lower GCS scores indicate a higher probability of mortality
Richmond RAS Scale	<ul style="list-style-type: none"> Strong negative correlation Lower scores indicate a higher probability of mortality.
Secondary Diagnosis	<ul style="list-style-type: none"> Very weak positive correlation.
Heart Rate and Respiratory Rate	<ul style="list-style-type: none"> Weak positive correlation. Higher heart rate and respiratory rate may indicate an increased risk of mortality.
Blood Pressure	<ul style="list-style-type: none"> Almost no correlation.

EDA

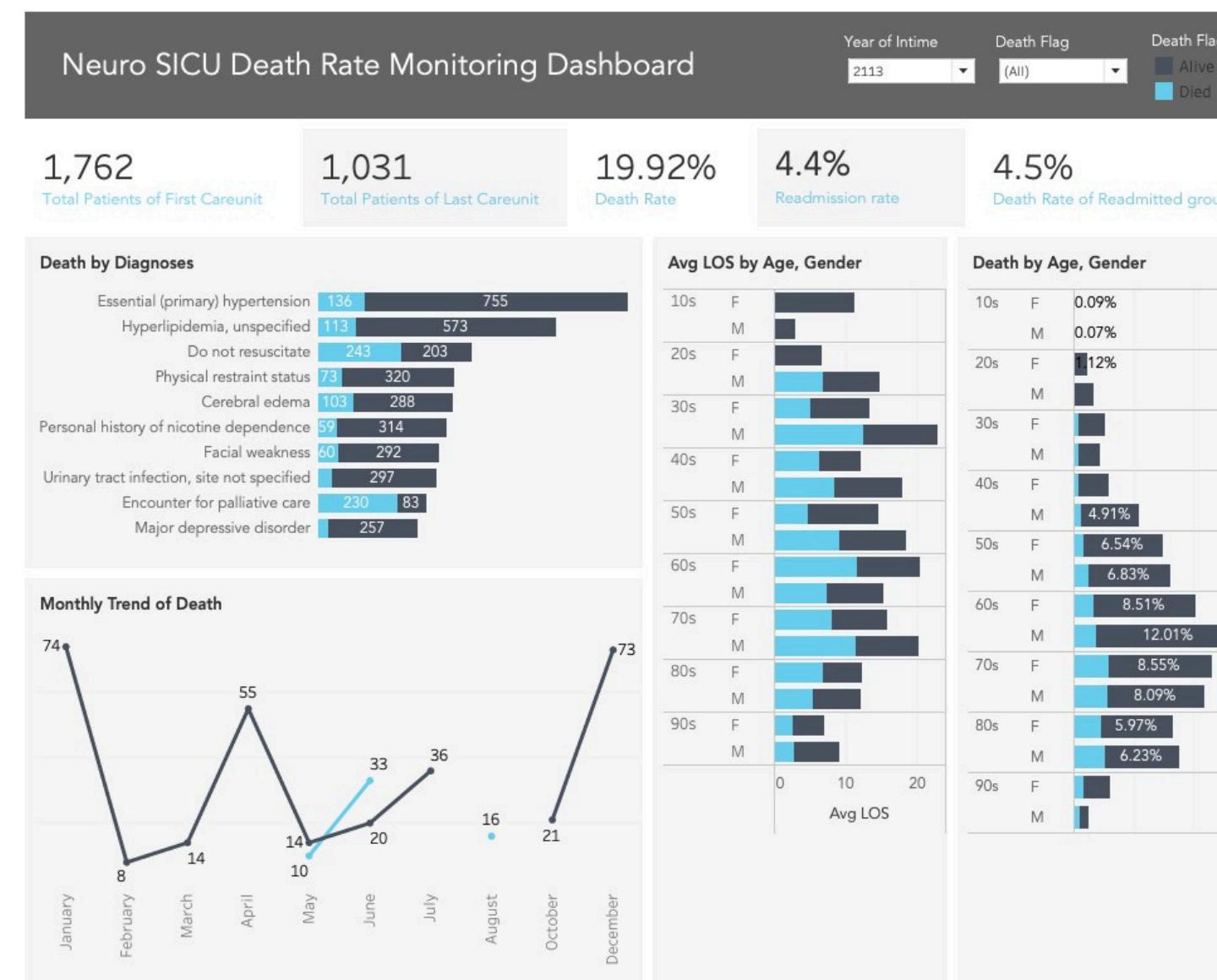
Correlation Analysis



Indicators of consciousness (GCS score, Richmond scale) and vital signs (heart rate, respiratory rate) have a stronger correlation with mortality rates

EDA

Tableau Dashboard Development



Objective

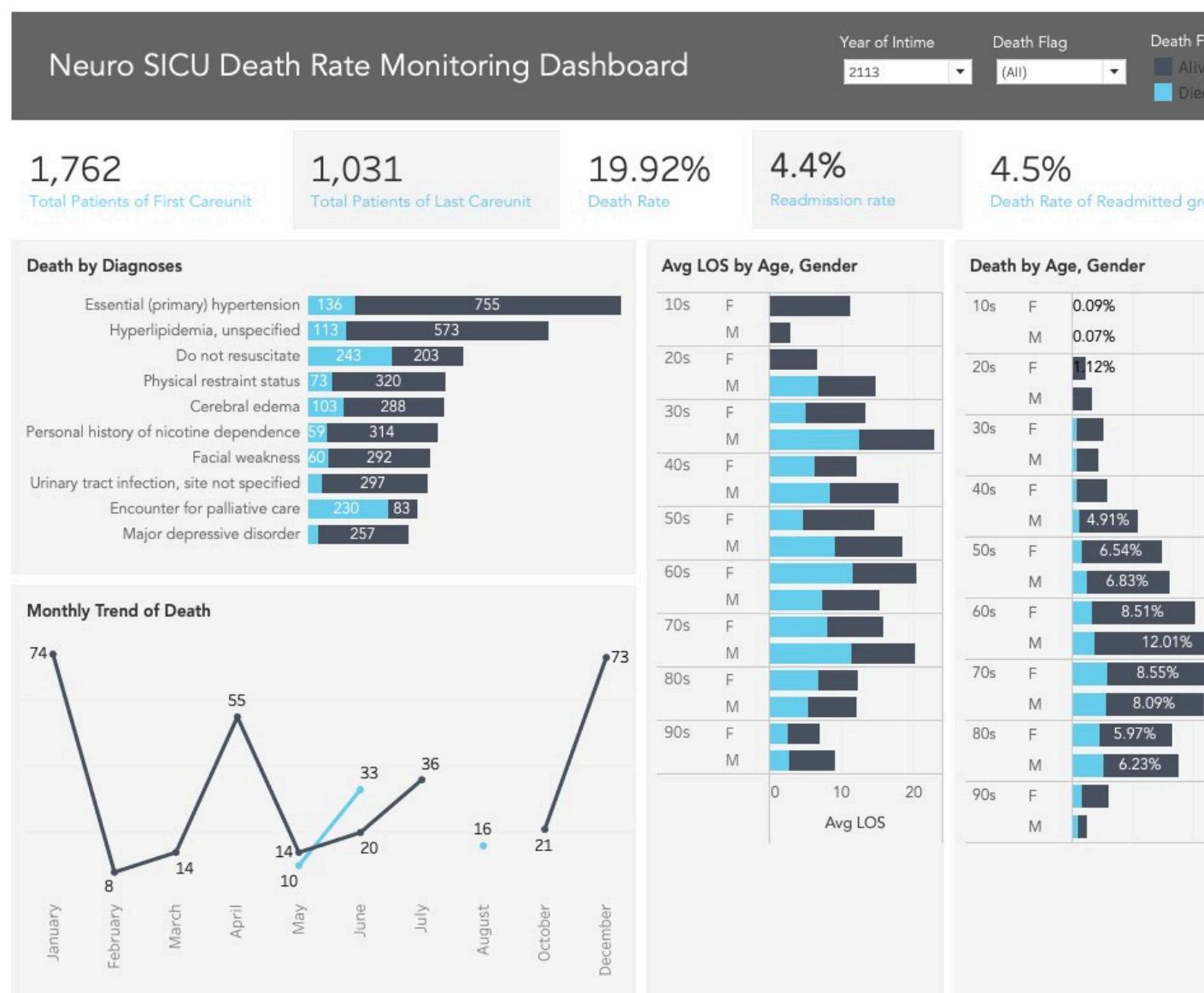
Monitoring the death rate in Neuro SICU

Type

KPI Dashboard

EDA

Tableau Dashboard Development



KPI

- Number of Patients by First and Last Care Unit
- Mortality and Readmission Rates
- Mortality Rate of the Readmission Group
- Number of Patients and Mortality Rate for Each Top 10 Diagnosis
- Monthly Mortality Trend
- Average Length of Stay by Age Group and Gender
- Mortality Rate by Age Group and Gender

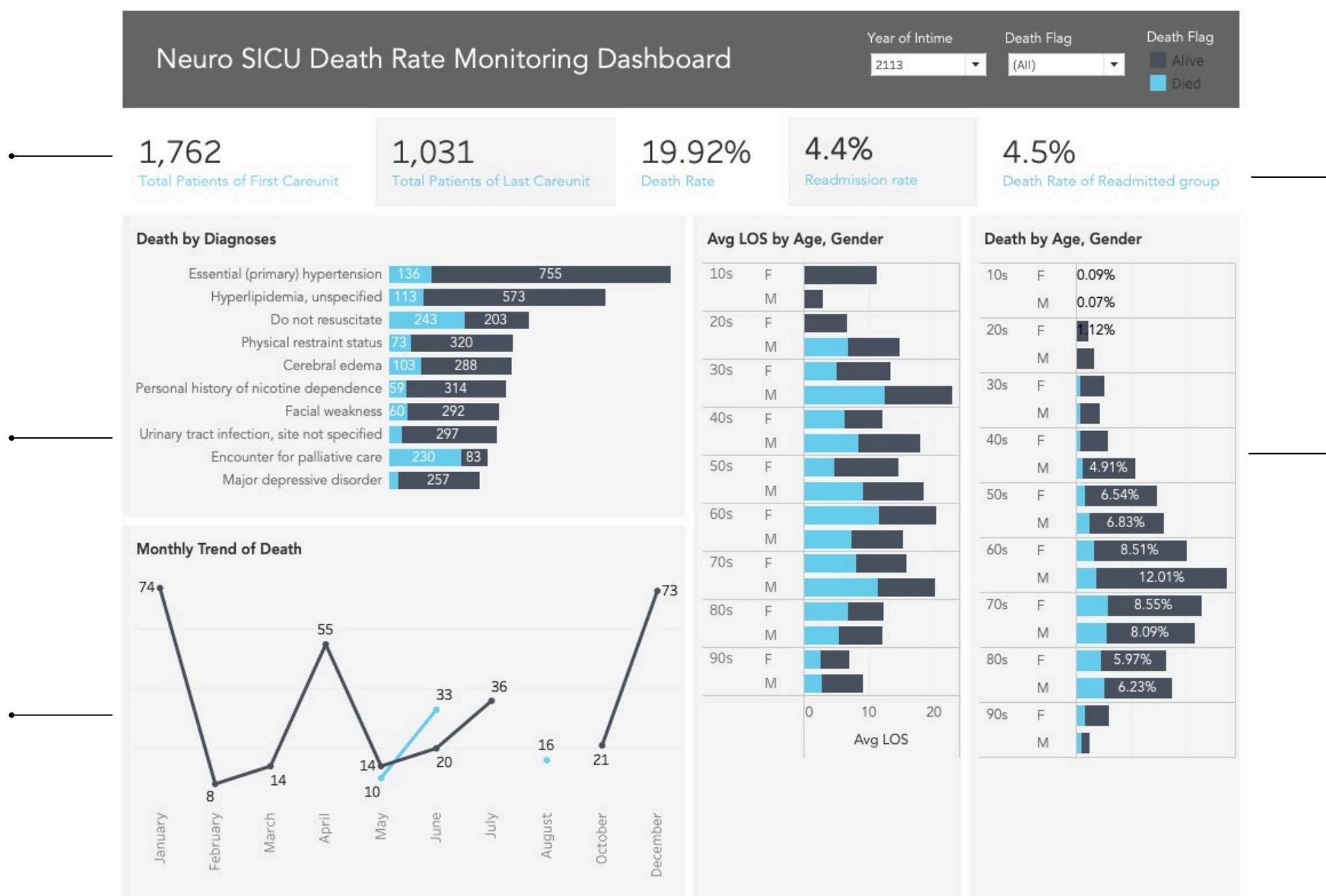
EDA

Tableau Dashboard Development

Identify congestion in specific units and find ways to **optimise transfer routes**

Improve treatment strategies for **specific diseases**

Reveal **seasonal factors** or specific events that affect mortality



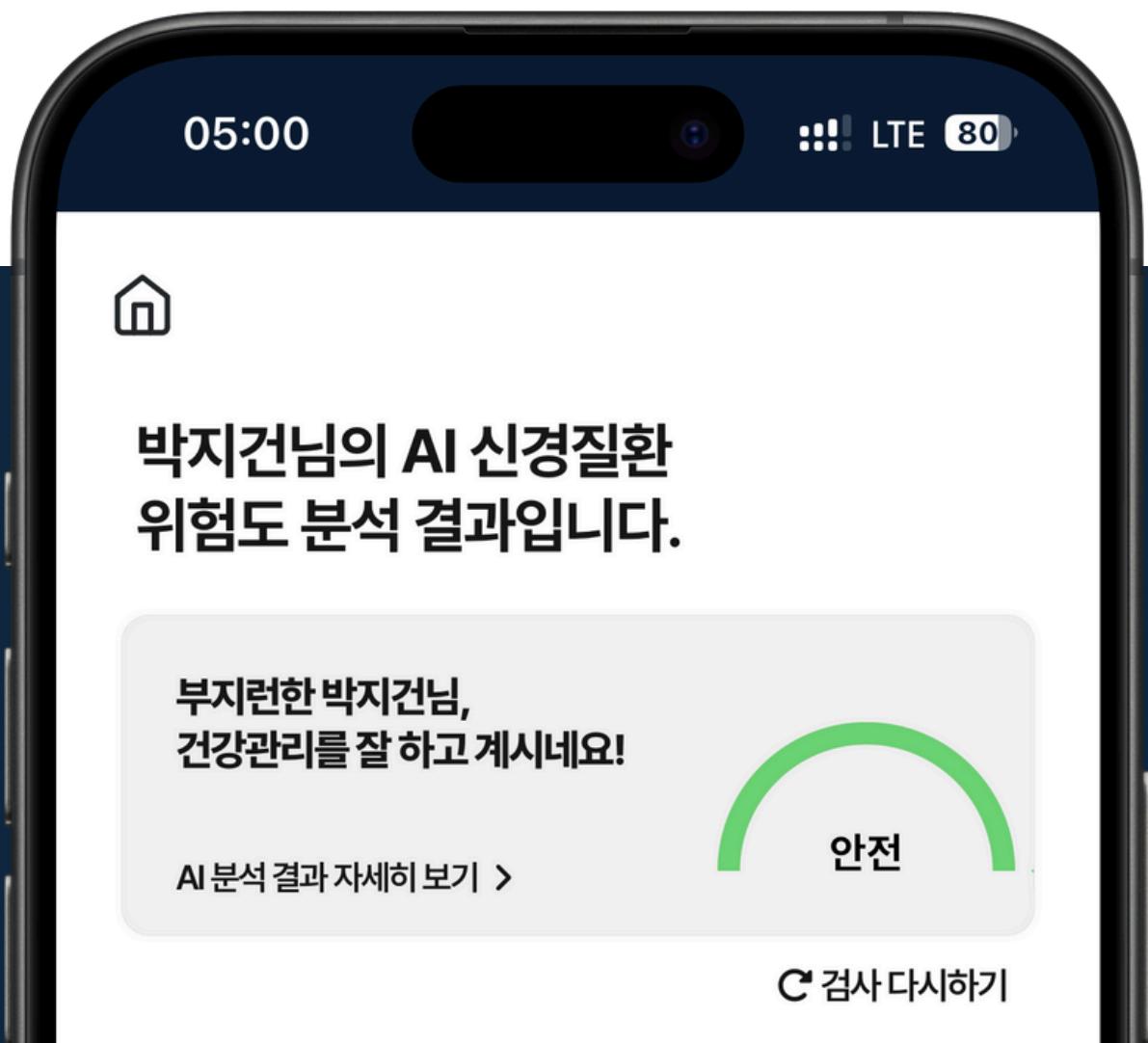
Evaluate the overall performance of the hospital by analysing the mortality and readmission rates

Optimise healthcare service delivery for high-risk groups



Machine Learning

Preprocessing / Modeling / Evaluation / Interpretation



ML Dept

Youjin Hwang (ML)

Data Preprocessing

Variable Selection

Modeling

Evaluation

Interpretation

**User friendly
Model**

**Precise
Algorithm**



Data Preprocessing

Defining variable type

Modeling

Evaluation

Interpretation

Data Preprocessing

Problem

Categorical variables (e.g.,
secondary conditions) are
incorrectly entered as
numerical data

Reason

Assigning **size and proportion** to each numerical category

Result

This does not learn the actual structure of the data,
ignores **nonlinear relationships**
-> Applied normalization and standardization

Data Preprocessing

Defining variable type

Modeling

Evaluation

Interpretation

Data Preprocessing



Round according to each category

Data Preprocessing

Defining variable type

Modeling

Evaluation

Interpretation

Data Preprocessing



Convert to **categorical** variables

Data Preprocessing

Defining variable type

Modeling

Evaluation

Interpretation

Data Preprocessing



Identify columns for **label encoding and **one-hot** encoding.**

Data Preprocessing

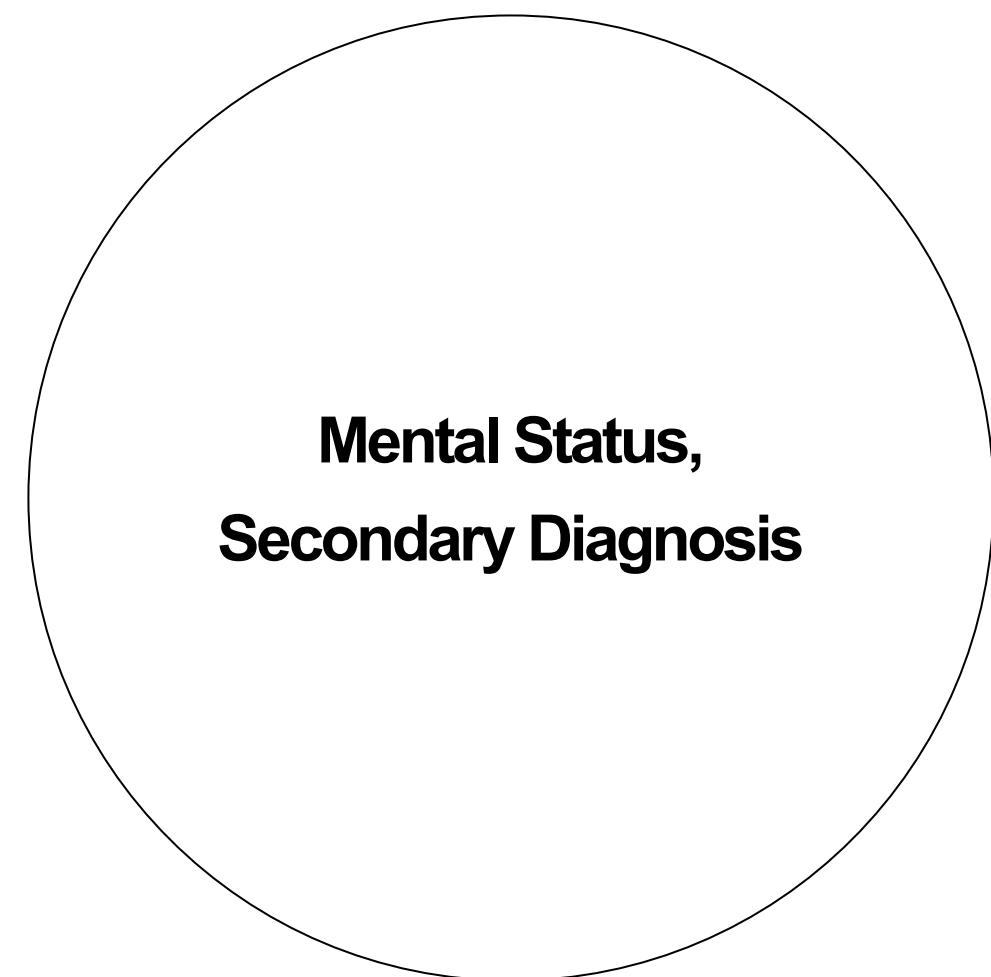
Defining variable type

Modeling

Evaluation

Interpretation

Data Preprocessing



Round according to **thresholds**

Data Preprocessing

Defining variable type

Modeling

Evaluation

Interpretation

Data Preprocessing



Convert to **categorical** variables

Data Preprocessing

Defining variable type

Modeling

Evaluation

Interpretation

Data Preprocessing



Add to columns for **one-hot encoding**.

Data Preprocessing

Scaling / Encoding

Modeling

Evaluation

Interpretation

Data Preprocessing

Numerical Variables

Standard Scaler

Categorical Variables

Label & One-Hot Encoding

Data Preprocessing

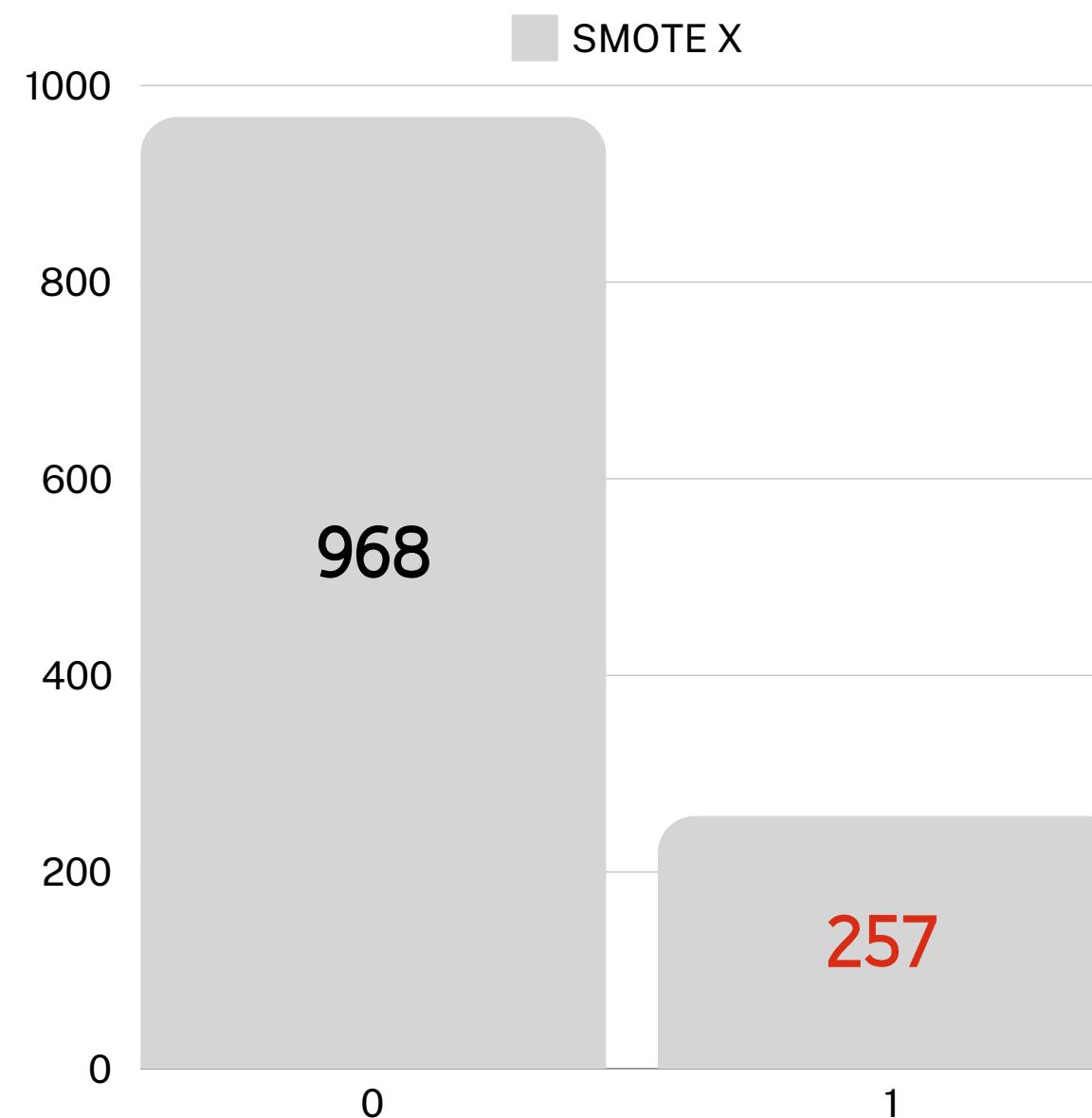
Resolve Class Imbalance

Modeling

Evaluation

Interpretation

Data Preprocessing



Data Preprocessing

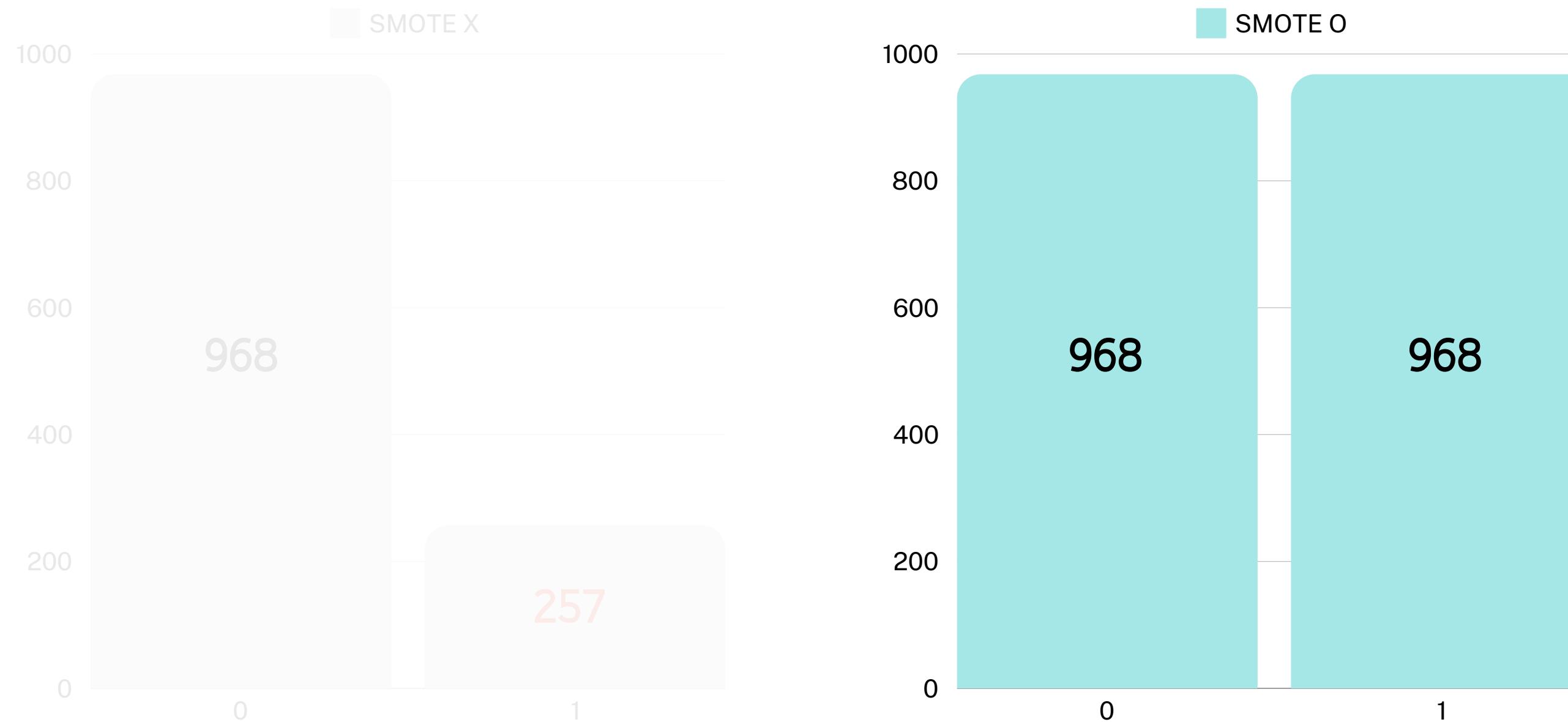
Resolve Class Imbalance

Modeling

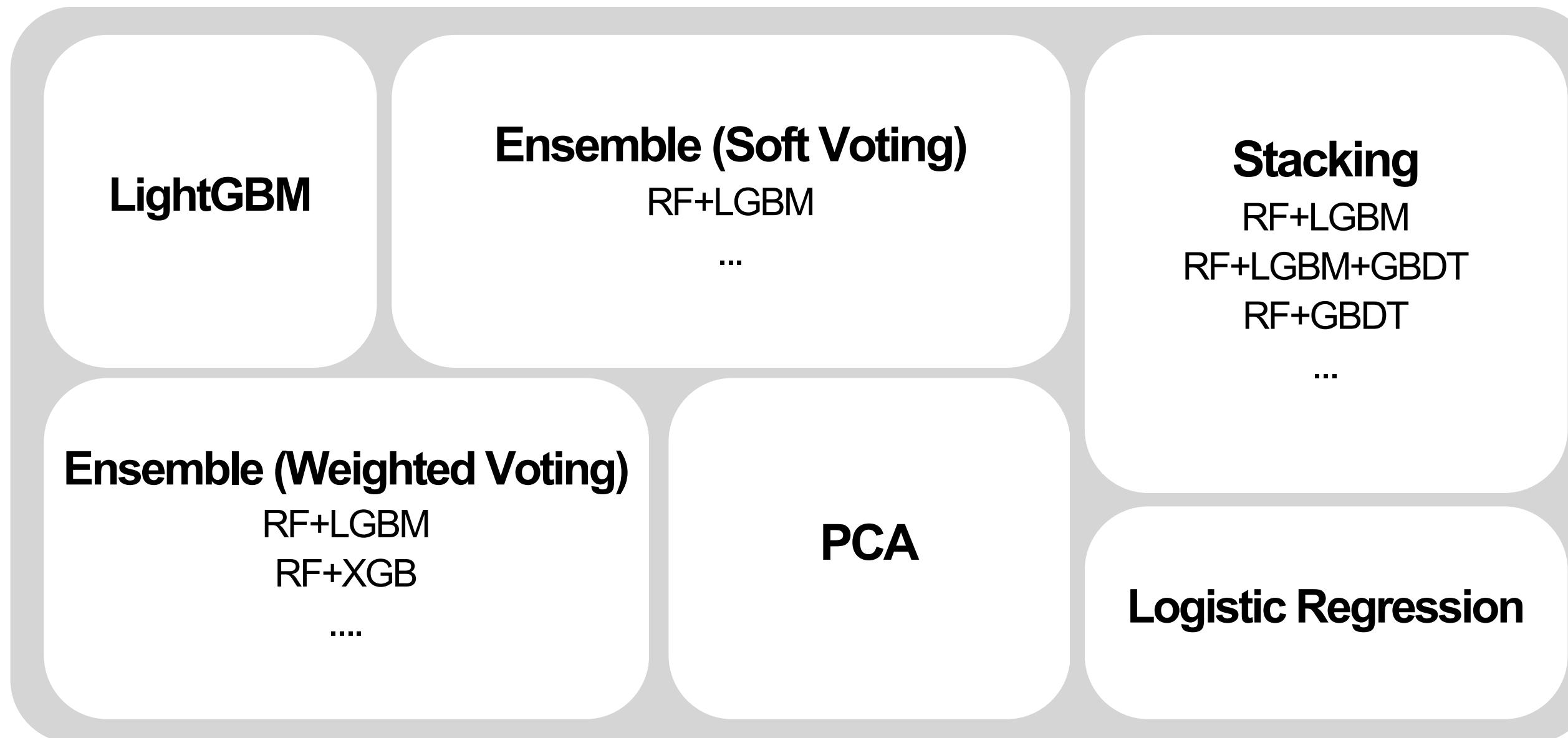
Evaluation

Interpretation

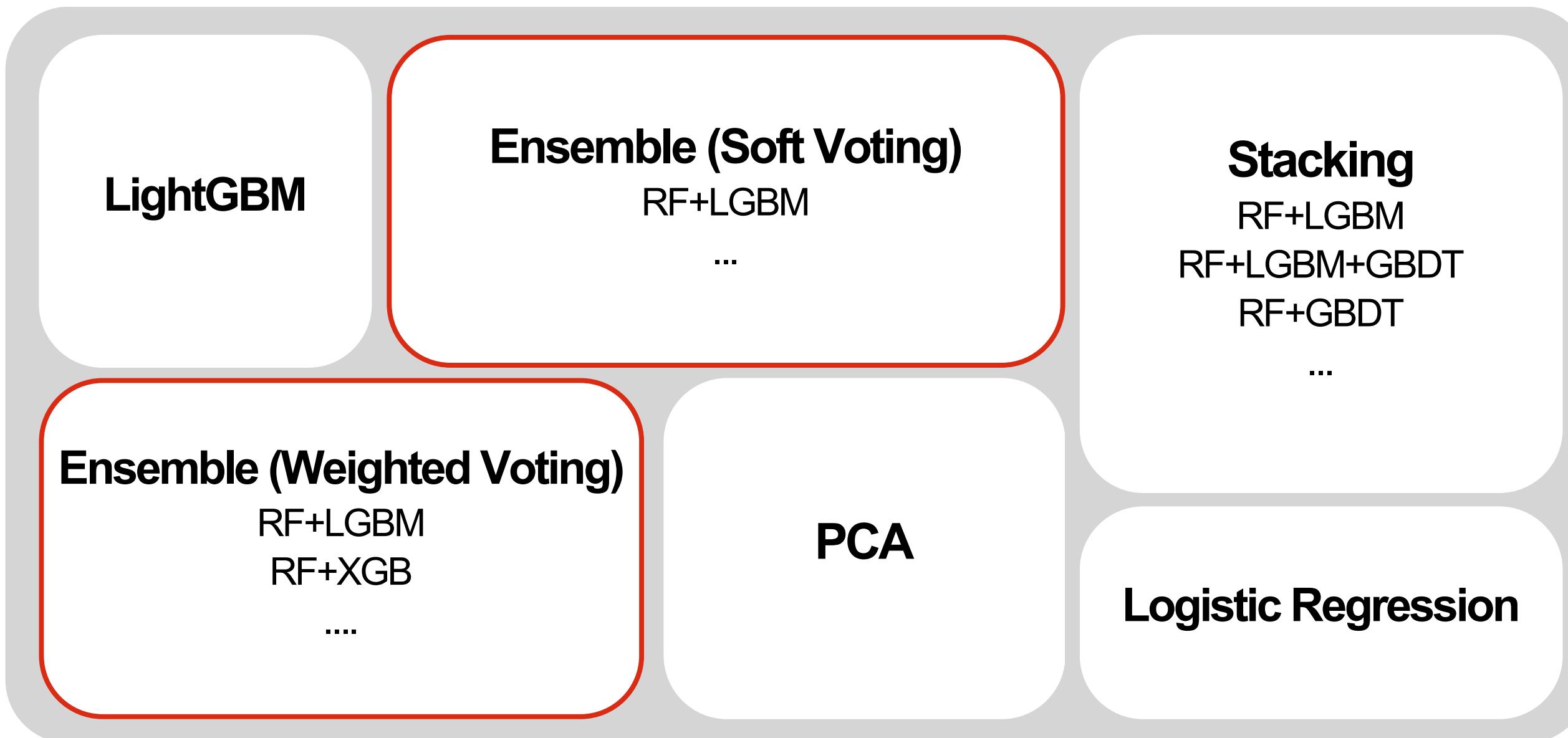
Data Preprocessing



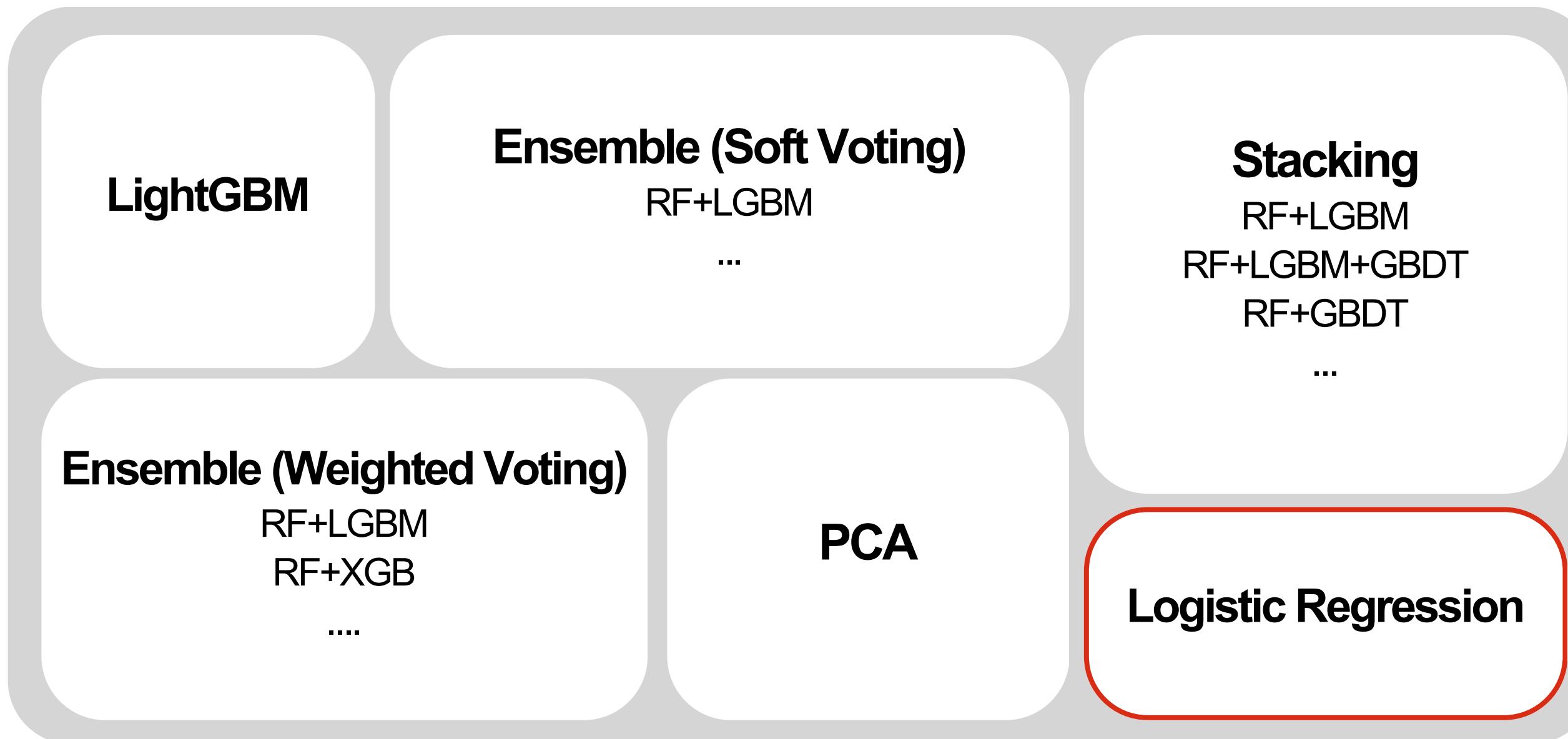
Modeling



Modeling

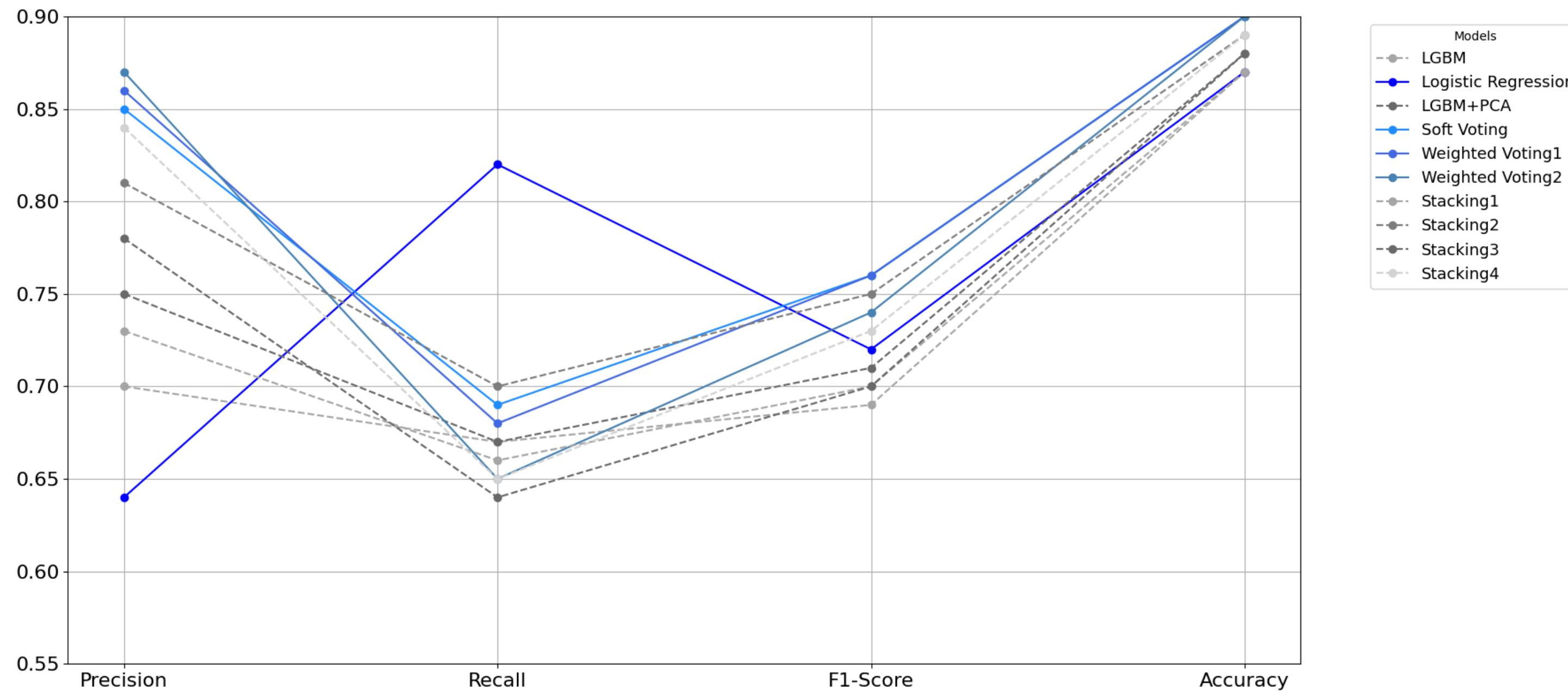


Modeling



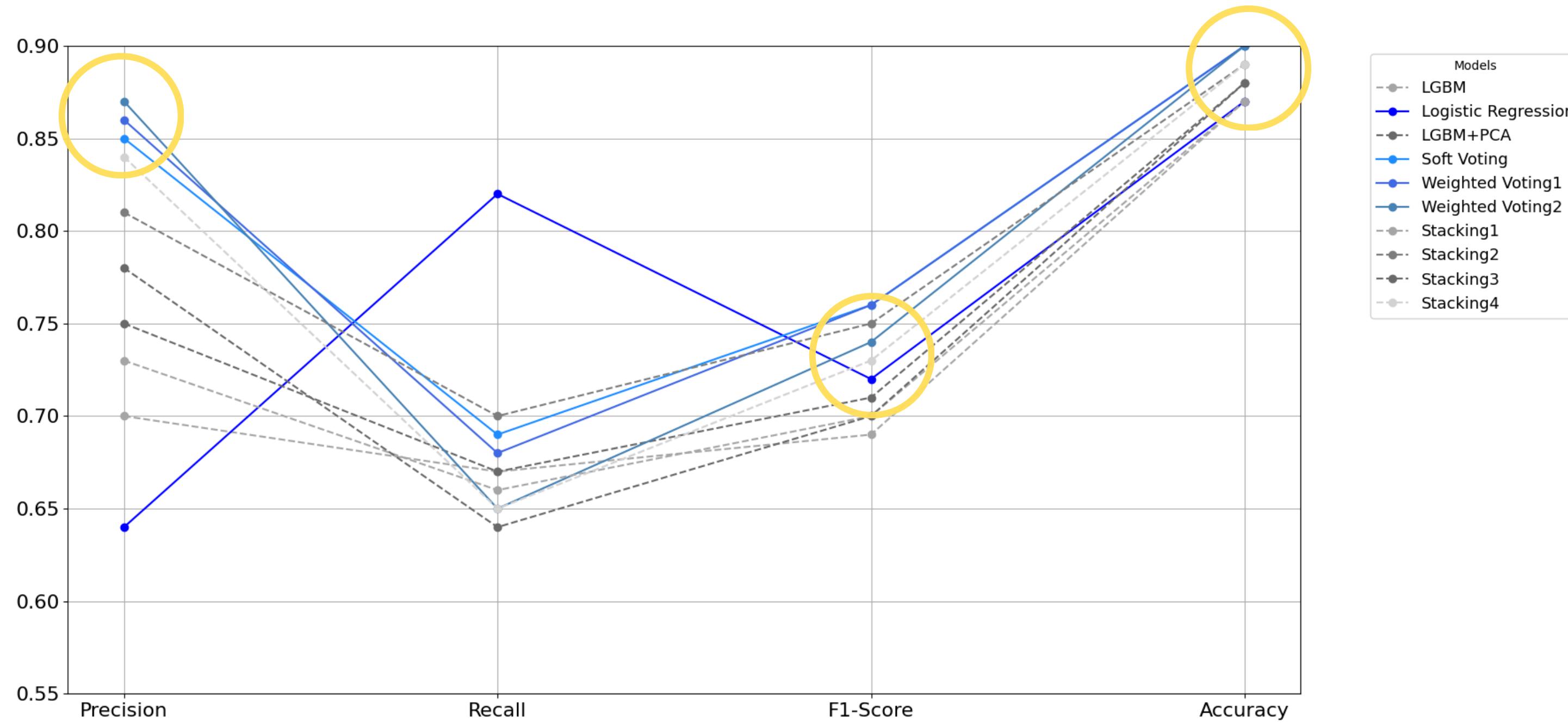
Comparison by Model

Evaluation



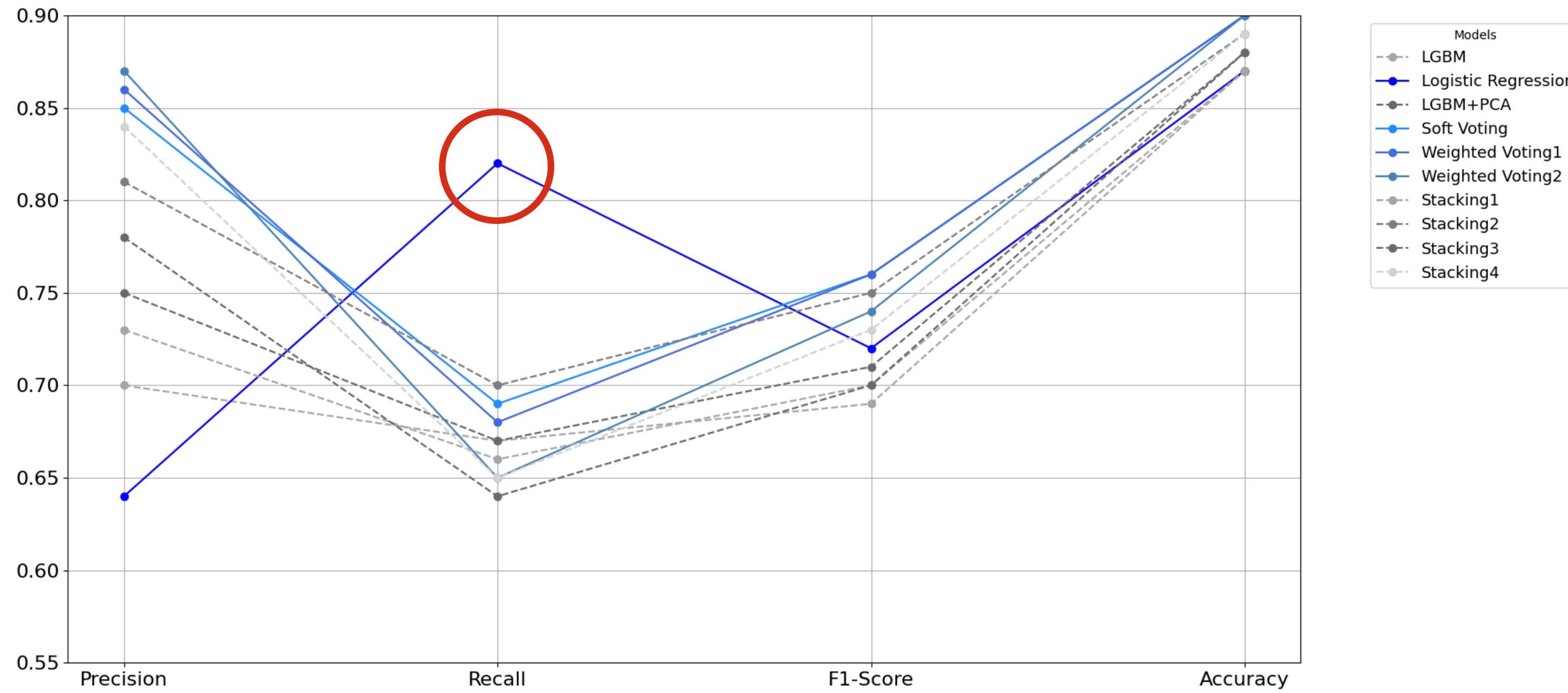
Comparison by Model

Evaluation



Comparison by Model

Evaluation



Data Preprocessing

Modeling

Evaluation

Comparison by Model

Interpretation

Logistic Regression

The **Highest Recall**

Ensemble

(**Soft Voting / Weighted Voting**)

The **Highest Precision, F1-Score,
Accuracy**

So, what **problems should we focus on?**

Data Preprocessing

Modeling

Evaluation

Comparison by Model

Interpretation

Predicted survival but resulted in death

Data Preprocessing

Modeling

Evaluation

Comparison by Model

Interpretation

Predicted death but survived

Data Preprocessing

Modeling

Evaluation

Comparison by Model

Interpretation

Logistic Regression

The **Highest Recall**

Ensemble

(Soft Voting / Weighted Voting)

The **Highest Precision, F1-Score,
Accuracy**

Data Preprocessing

Modeling

Evaluation

Comparison by Model

Interpretation

Logistic Regression

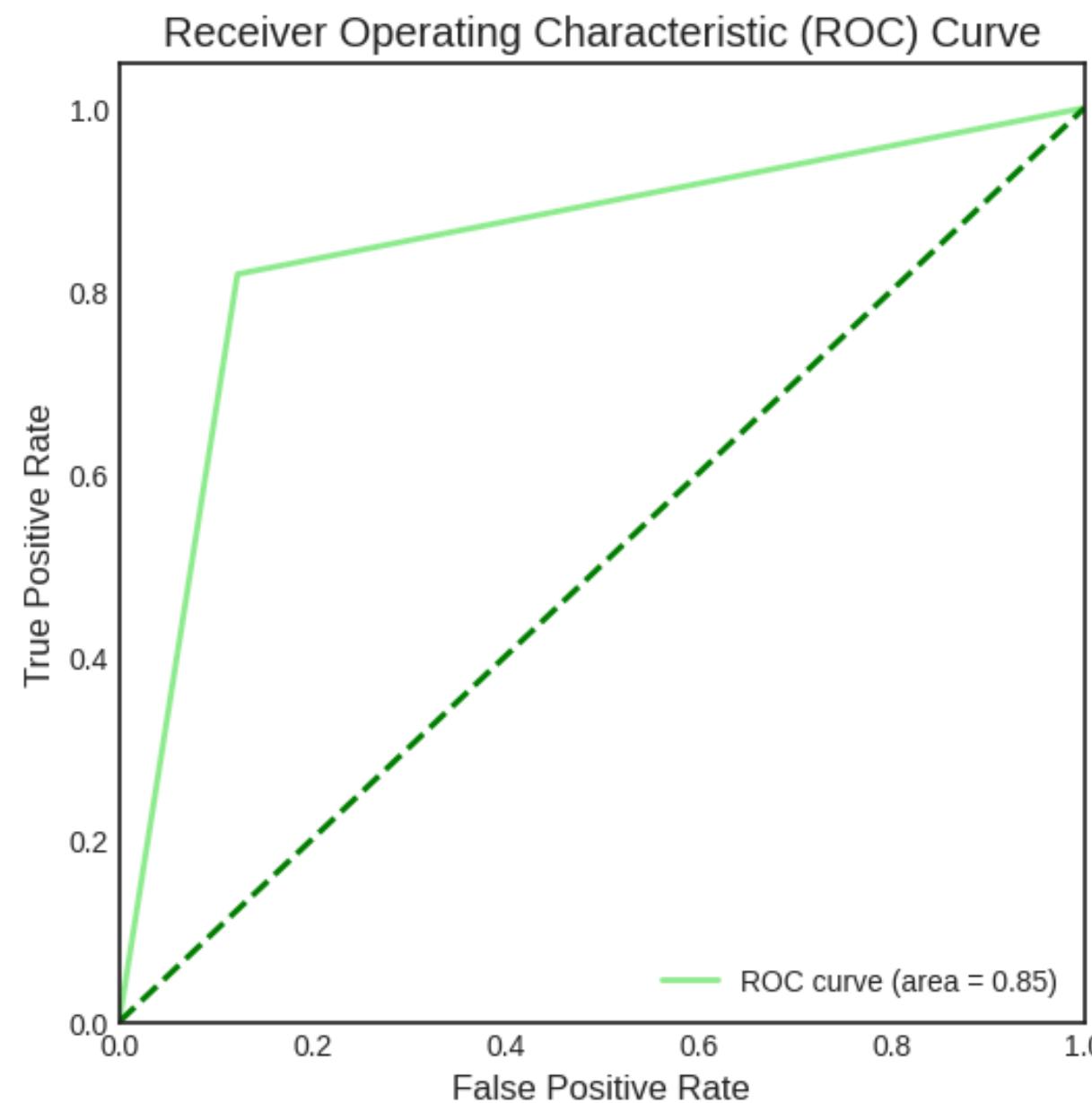
The Highest Recall

Ensemble

(Soft Voting / Weighted Voting)

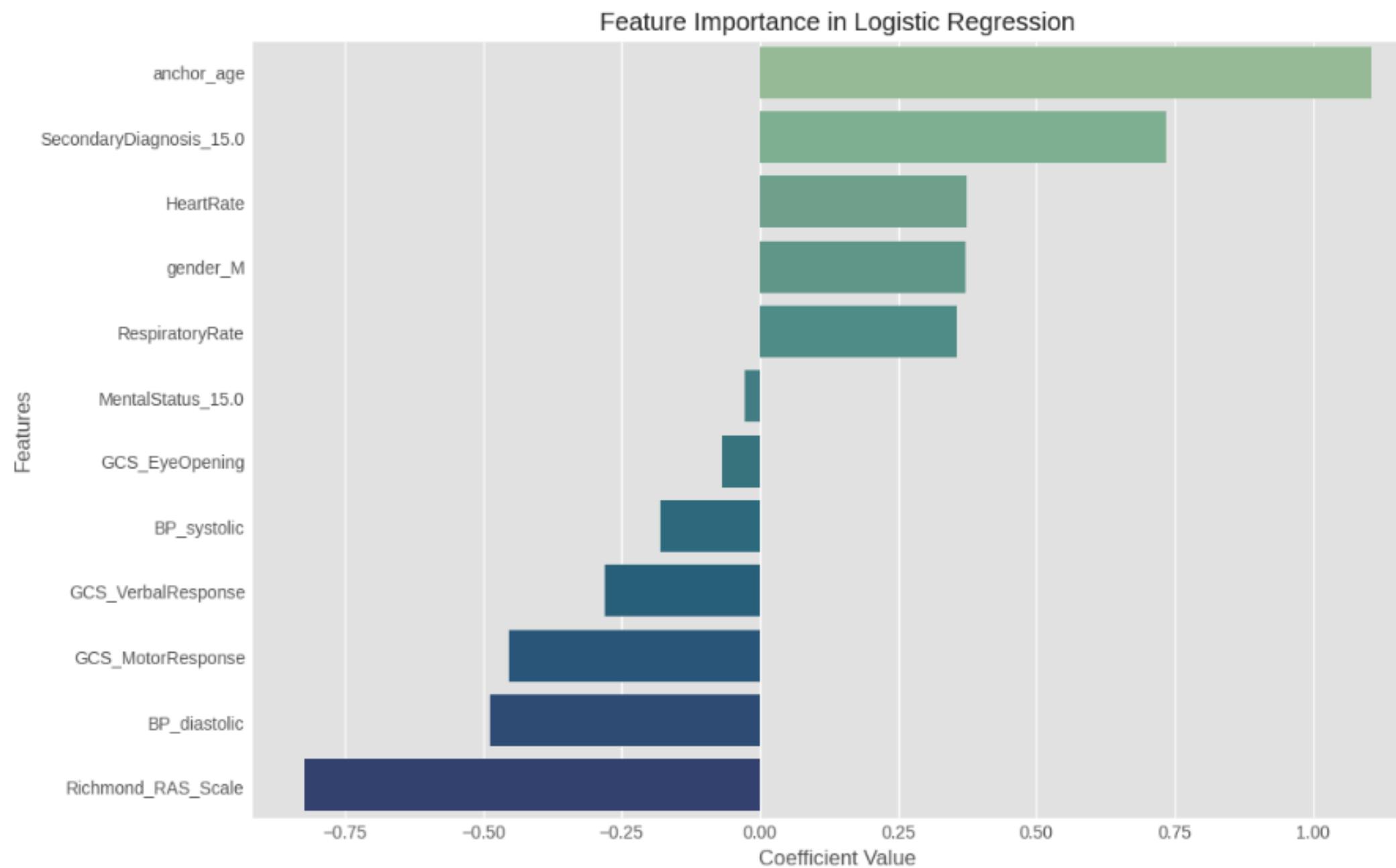
가장 높은 Precision, F1-Score,
Accuracy

Interpretation



AUROC : 0.85

Interpretation



+

Age, Secondary Diagnosis

-

RASS Scale, BP systolic,
GCS scores

Data Preprocessing

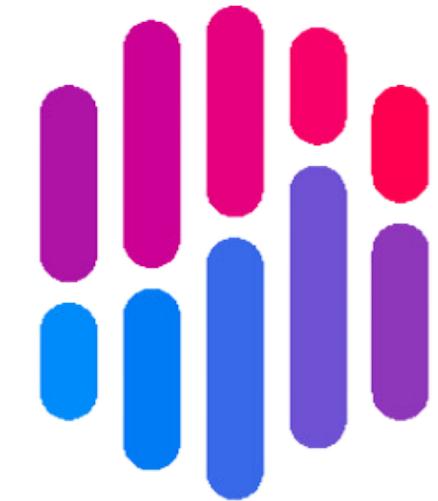
Modeling

Evaluation

Interpretation
XAI-SHAP

Interpretation

SHAP (SHapley Additive exPlanation)



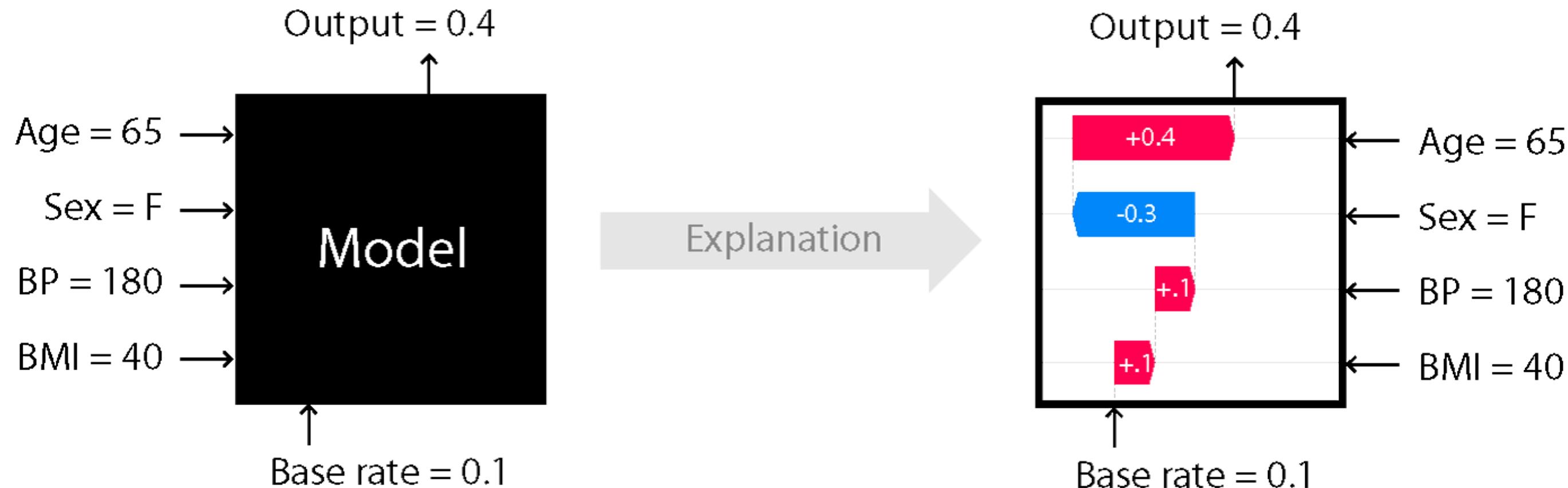
SHAP

eXplainable AI, XAI

: **SHAP** (SHapley Additive exPlanation)

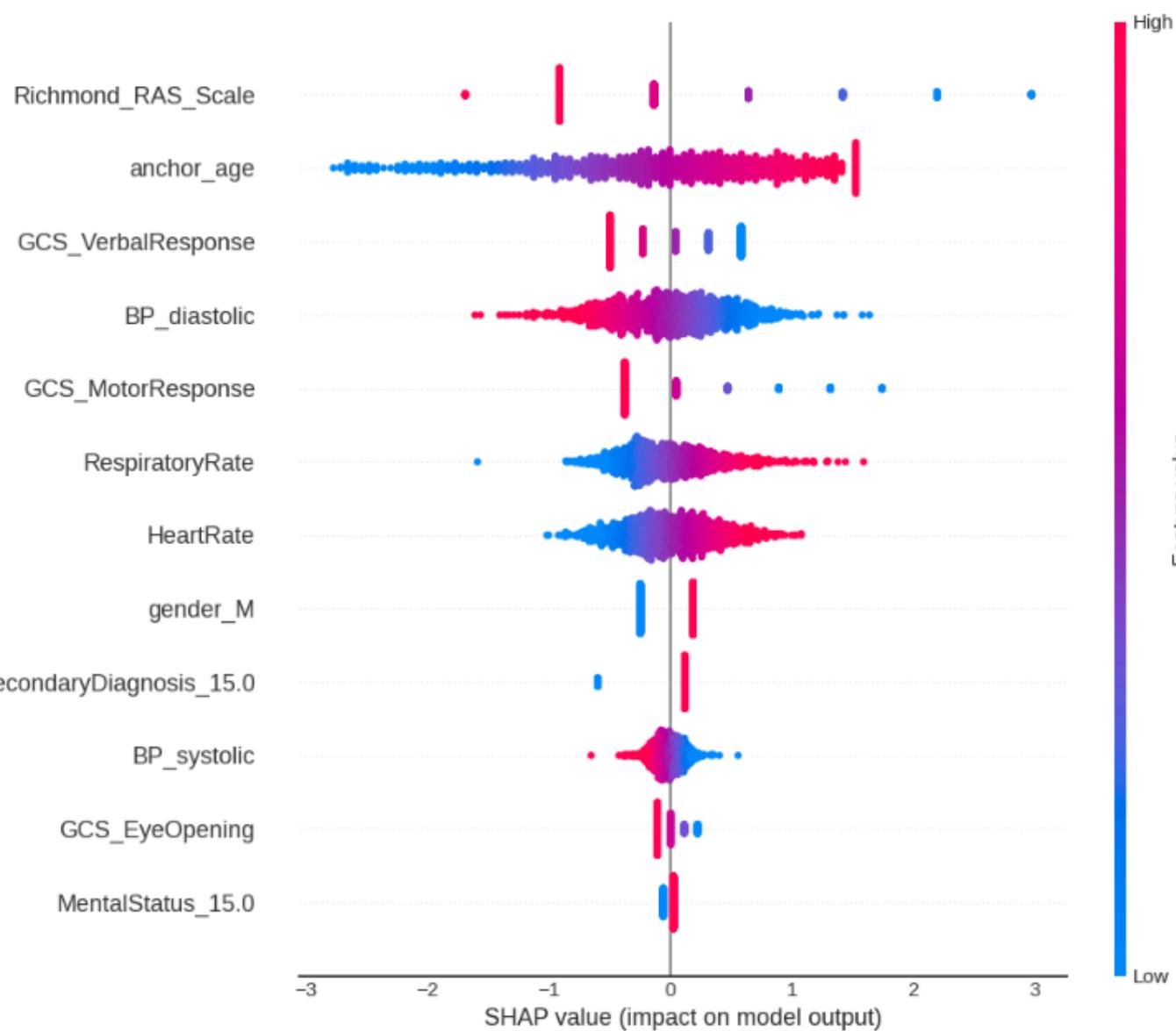
Interpretation

SHAP (SHapley Additive exPlanation)



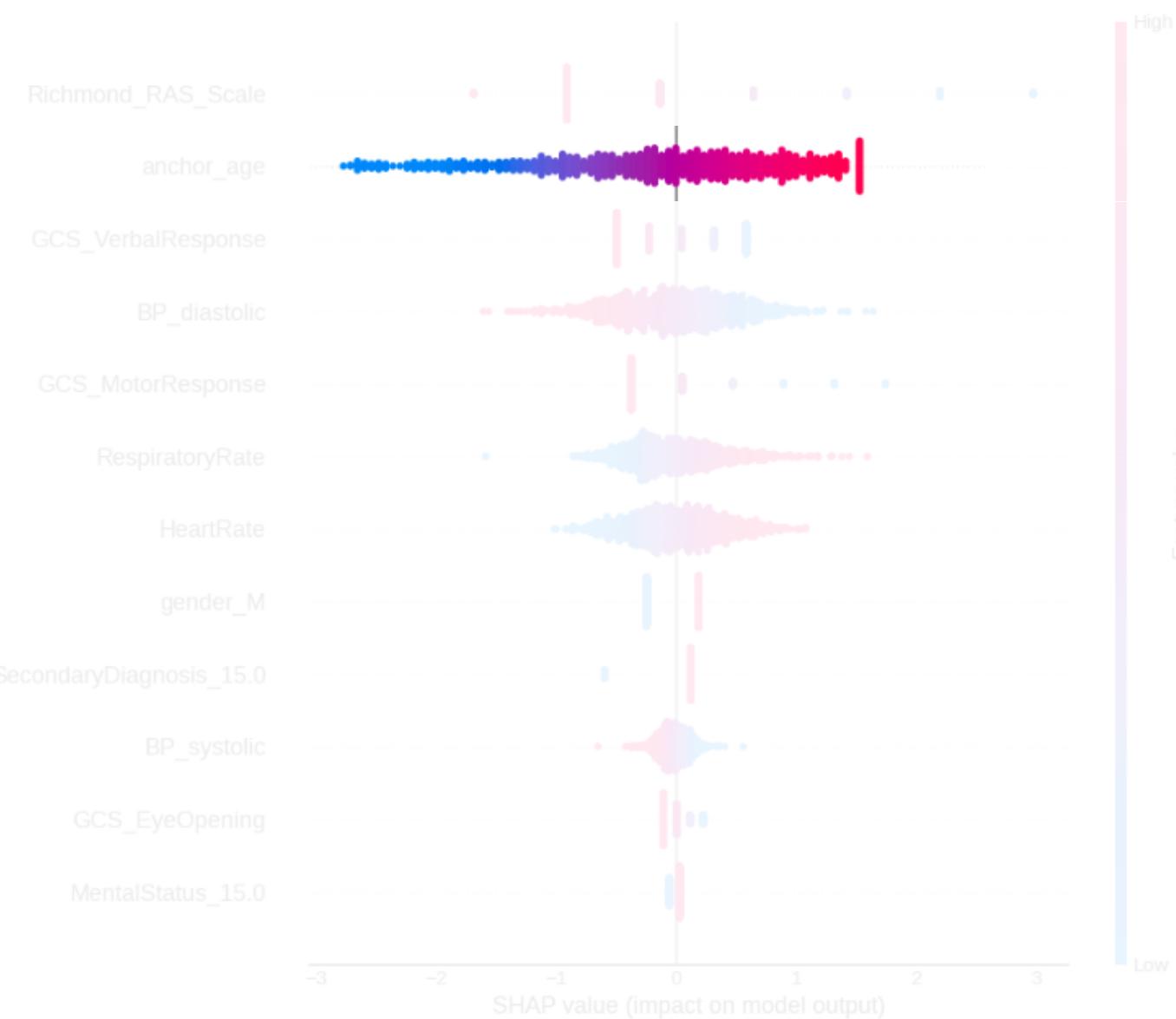
Interpretation

SHAP (SHapley Additive exPlanation)



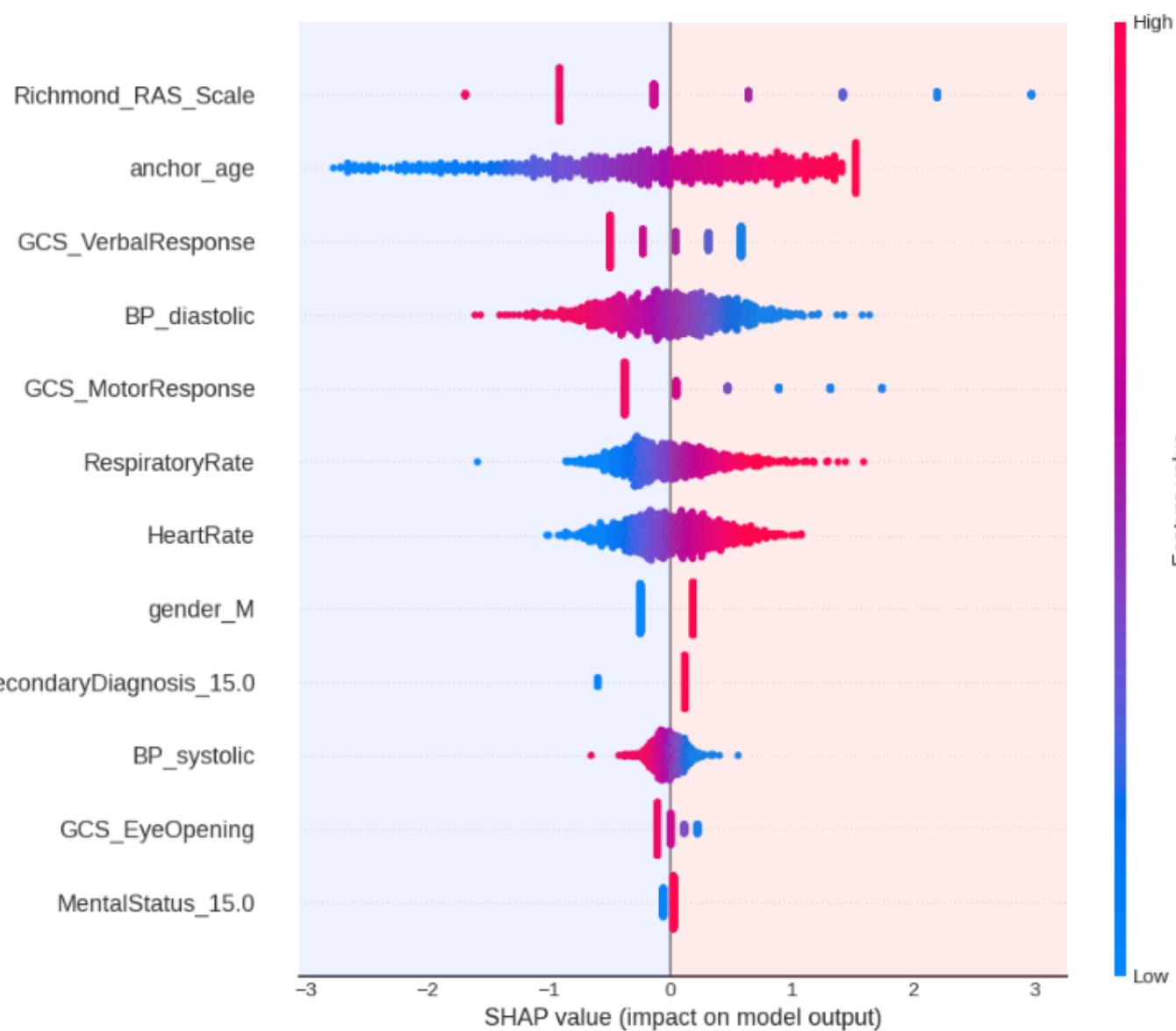
Interpretation

SHAP (SHapley Additive exPlanation)



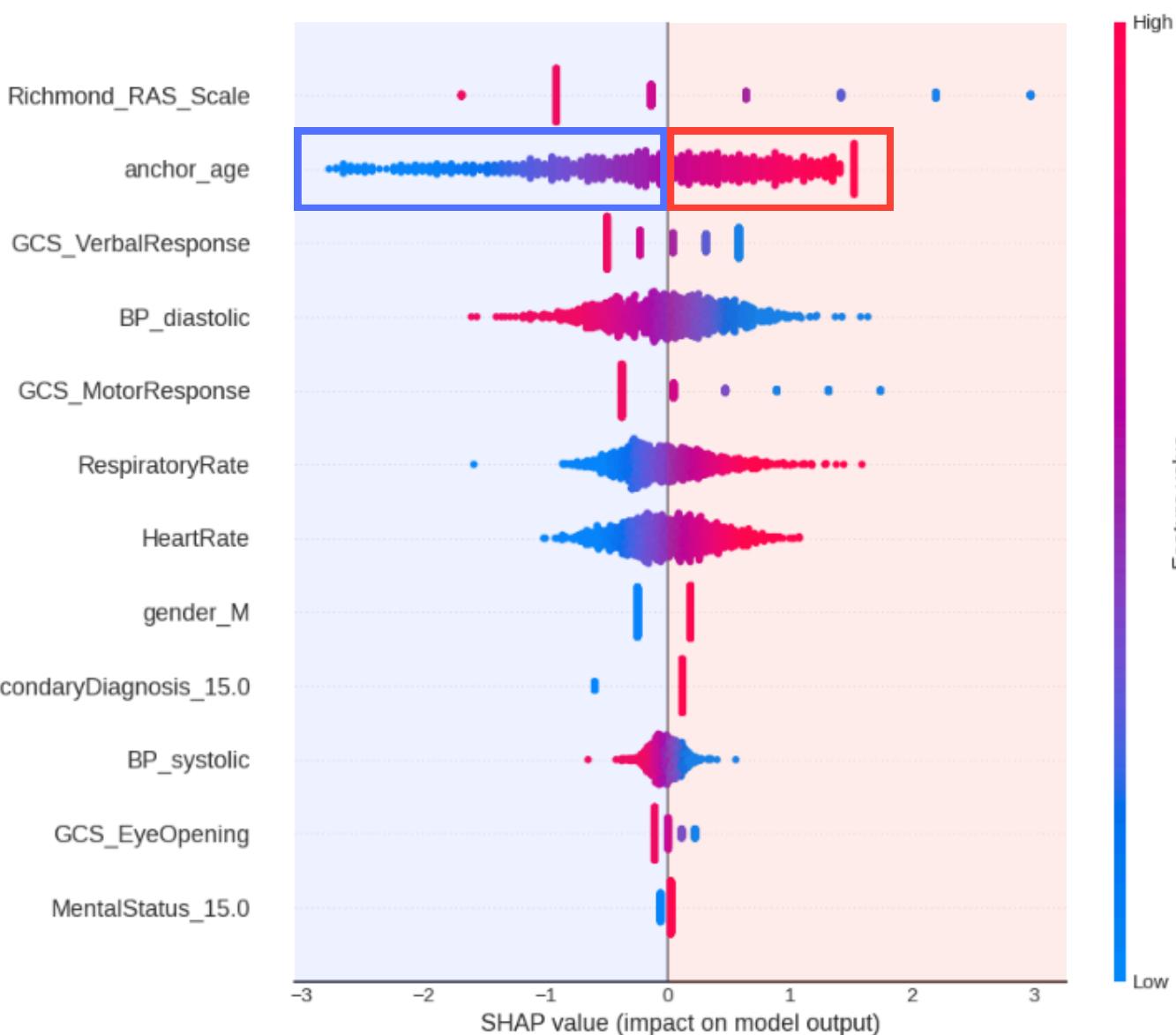
Interpretation

SHAP (SHapley Additive exPlanation)



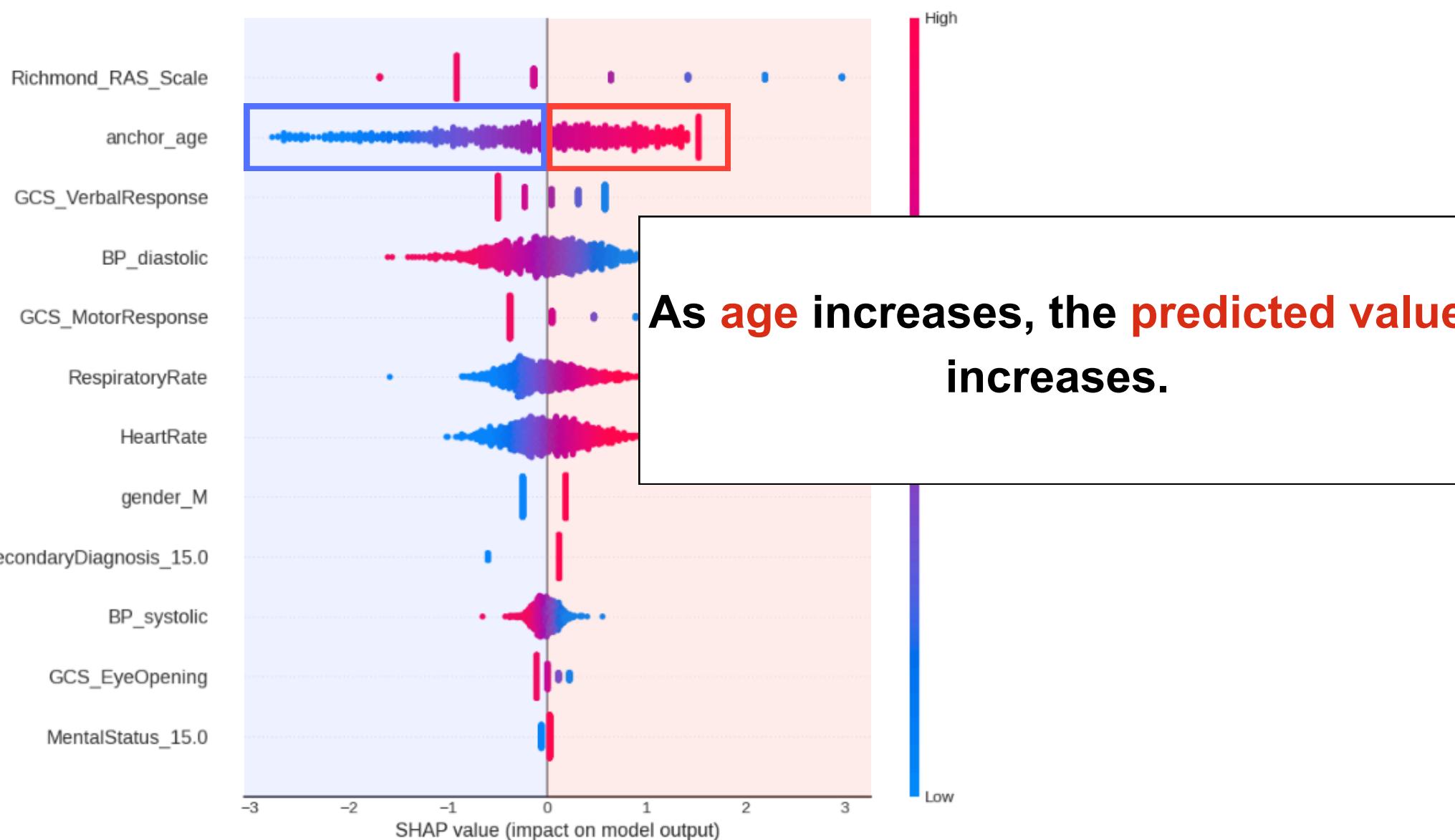
Interpretation

SHAP (SHapley Additive exPlanation)



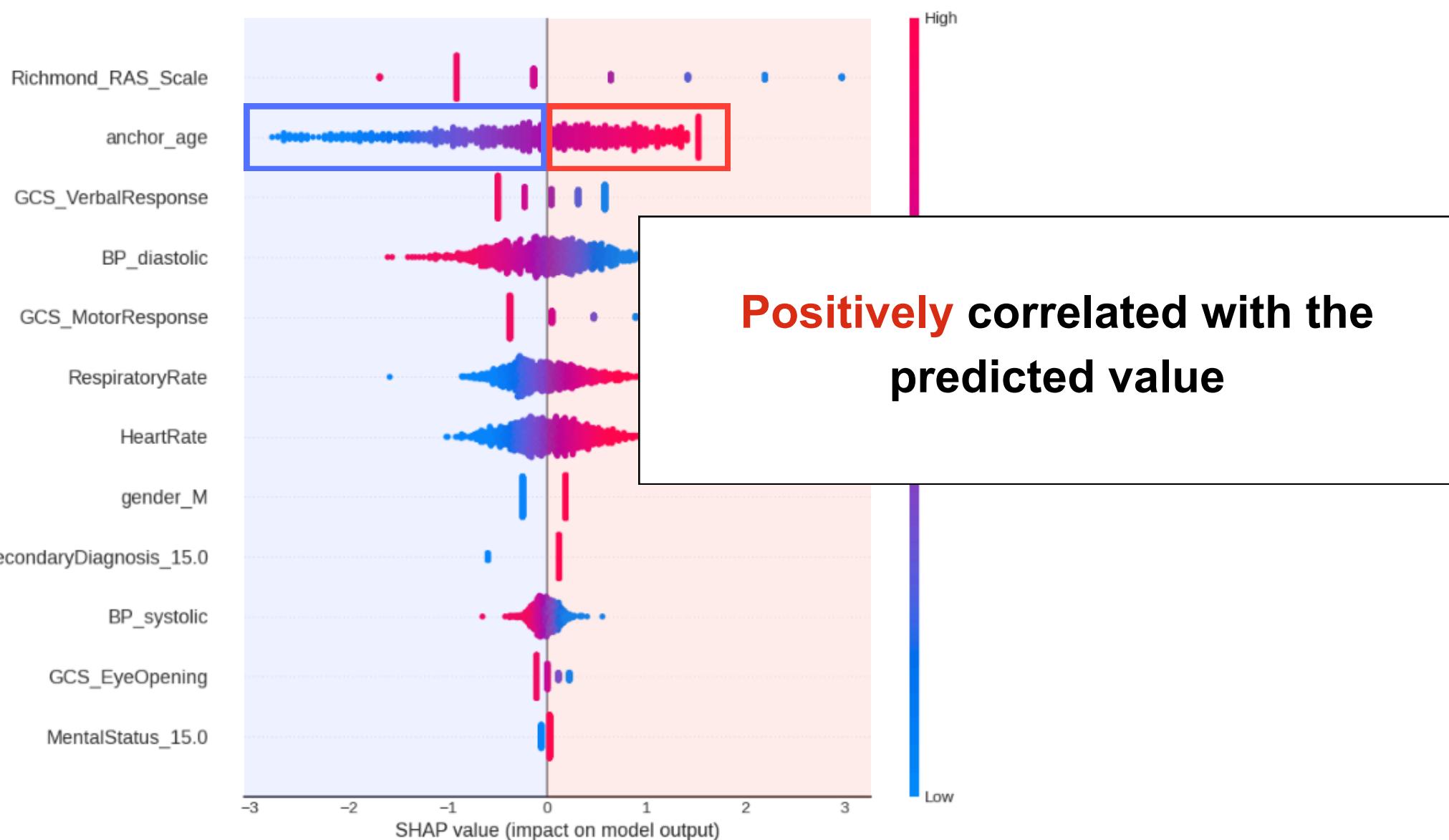
Interpretation

SHAP (SHapley Additive exPlanation)



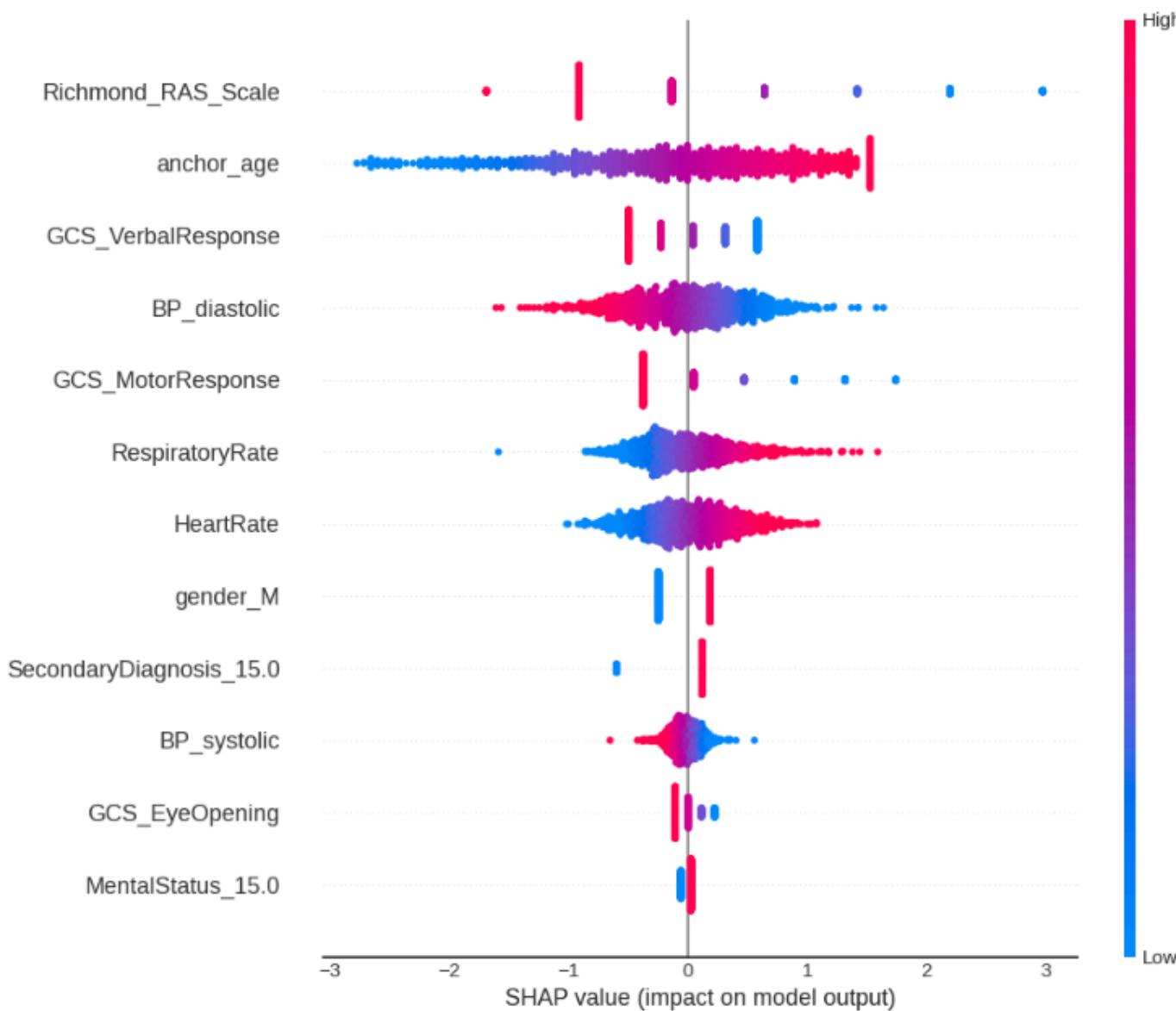
Interpretation

SHAP (SHapley Additive exPlanation)



Interpretation

SHAP (SHapley Additive exPlanation)

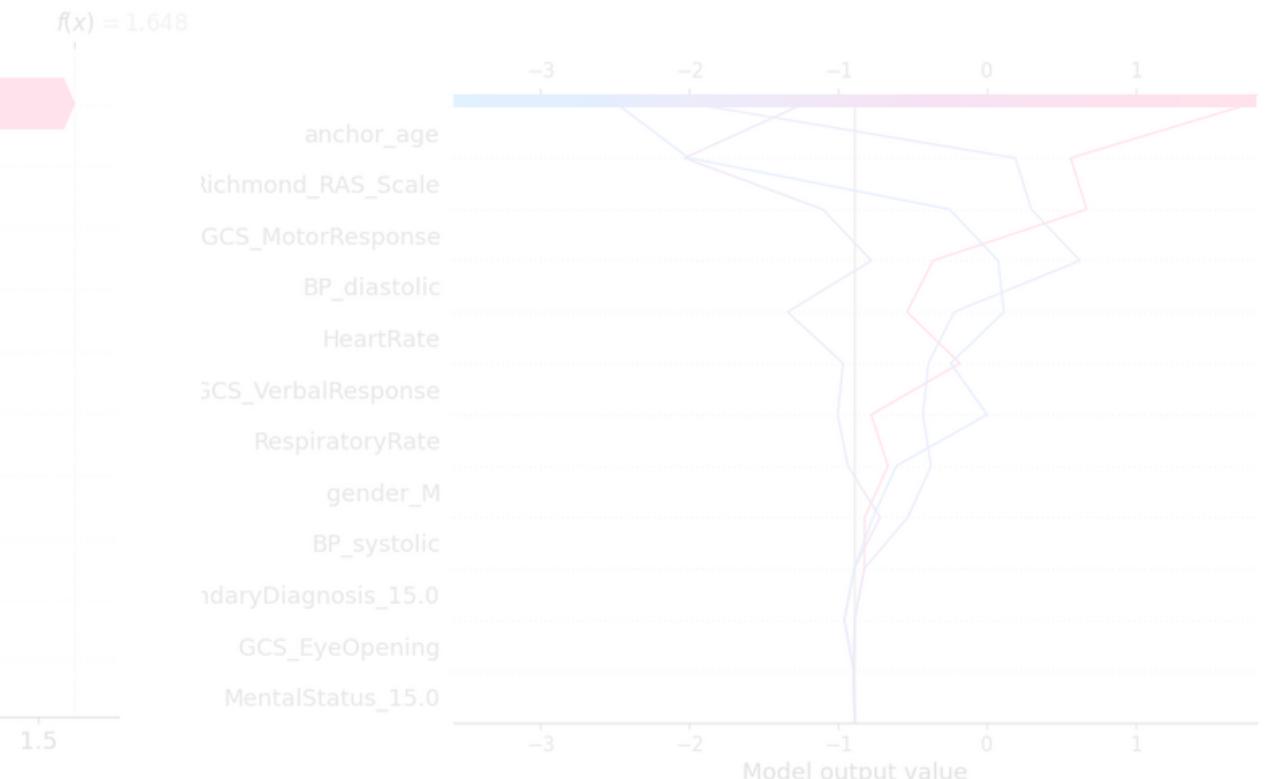
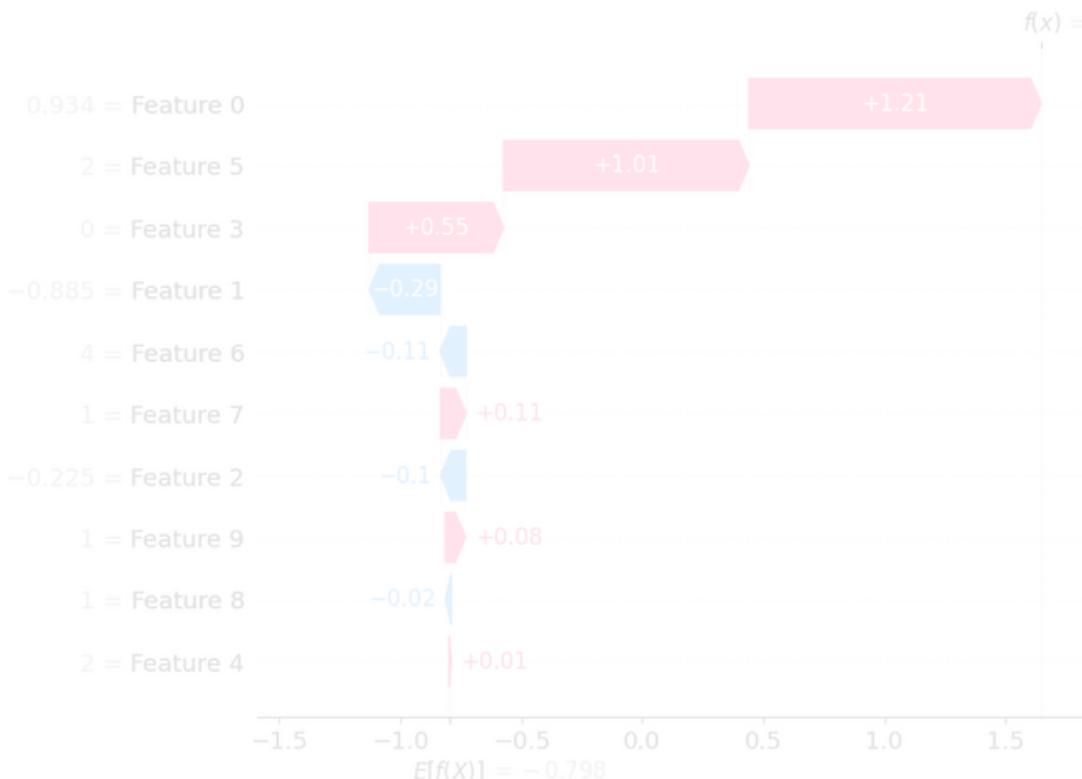
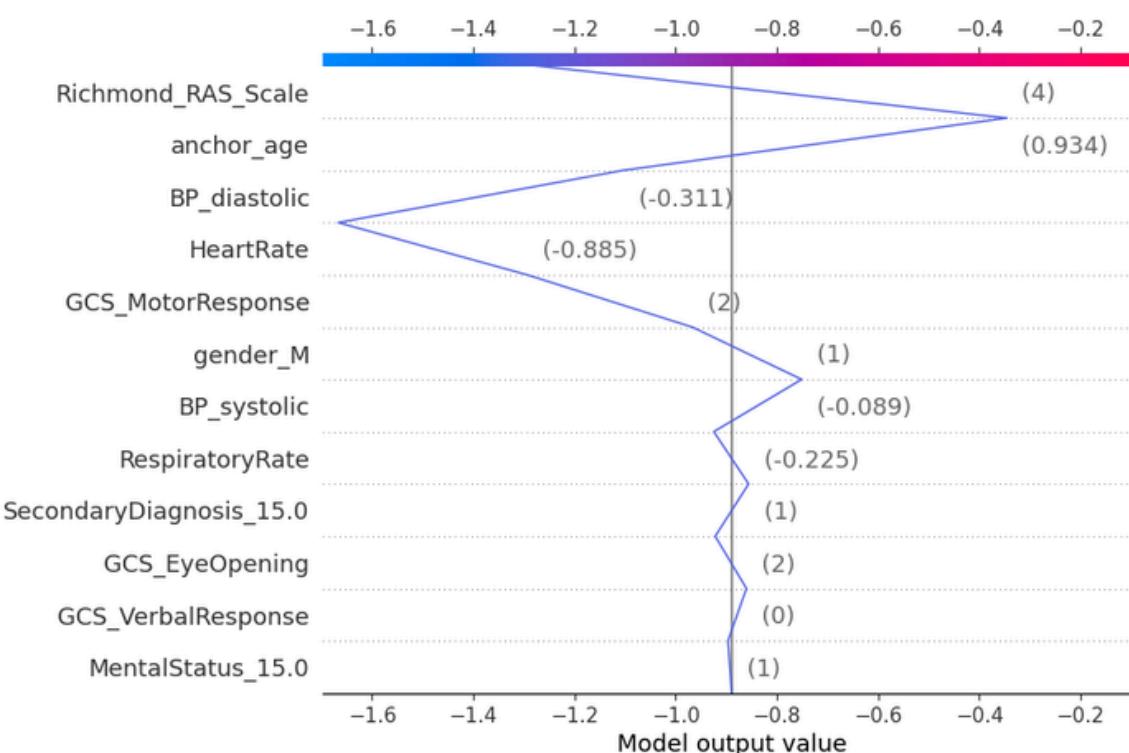


Positive(+) correlation
Age, Respiratory rate, Heart Rate, Gender(M),
Secondary diagnosis, Mental Status

Negative(-) correlation
RASS, GCS, Blood pressure

모델 해석

SHAP (SHapley Additive exPlanation)



Data Preprocessing

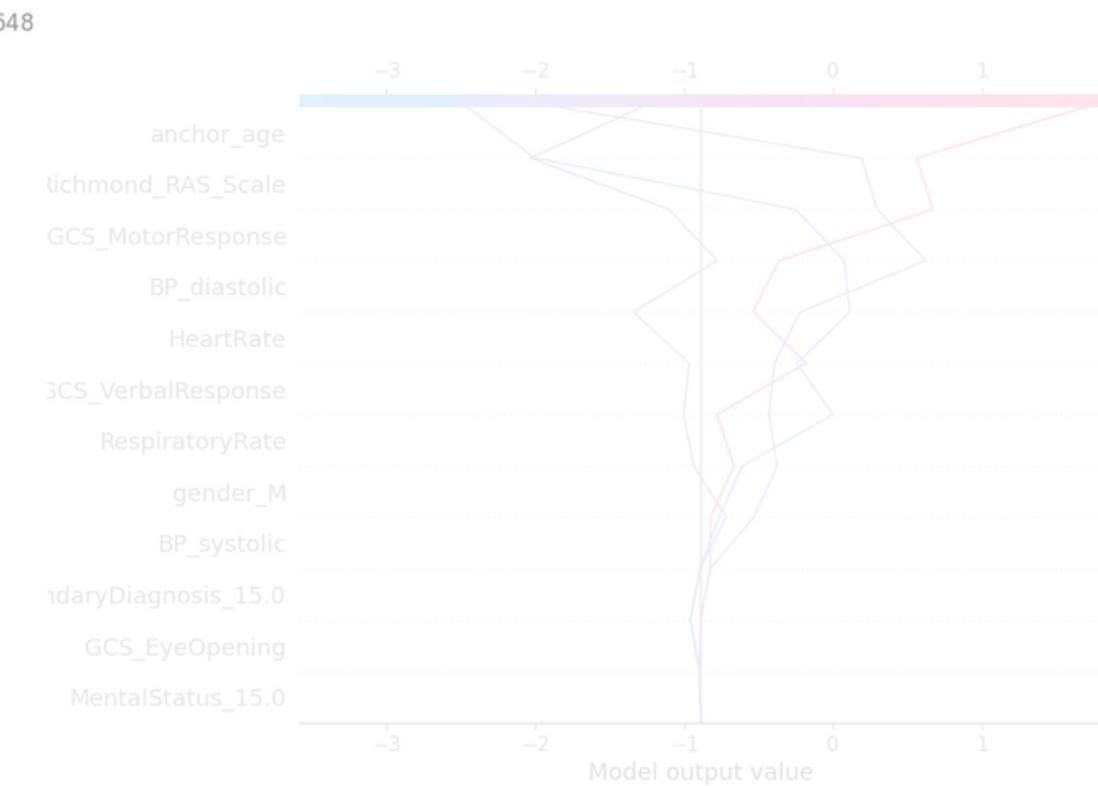
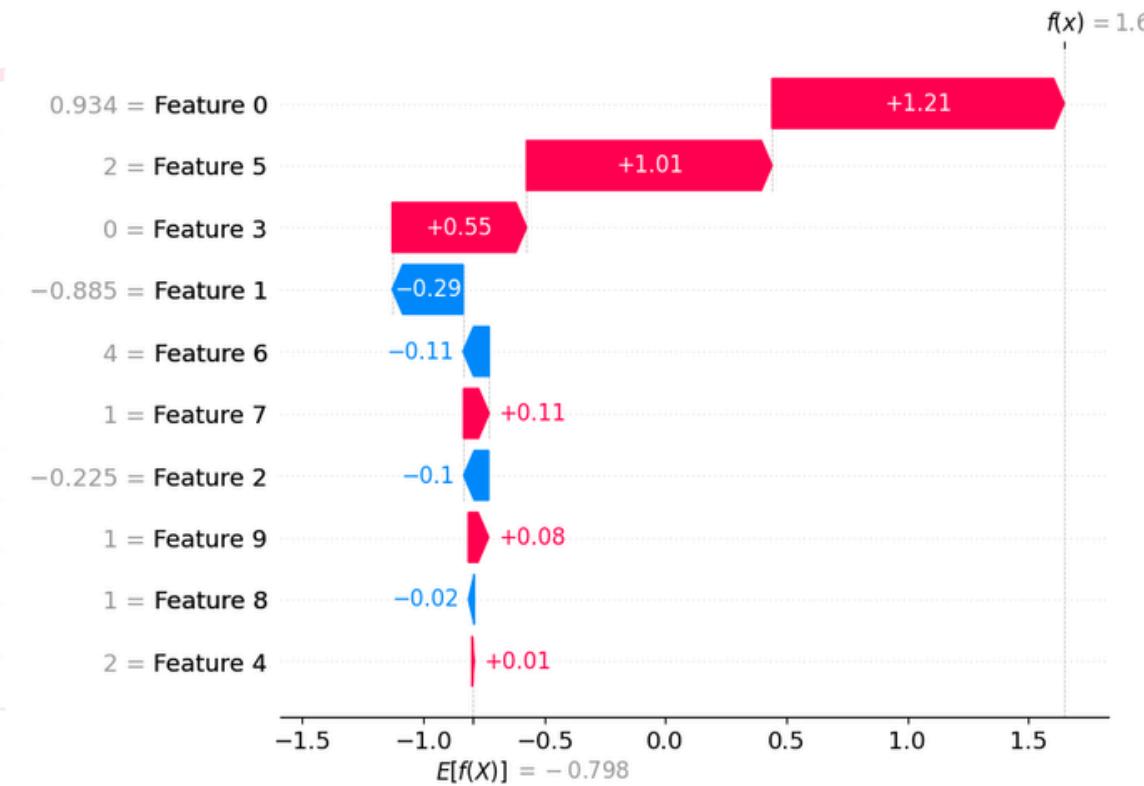
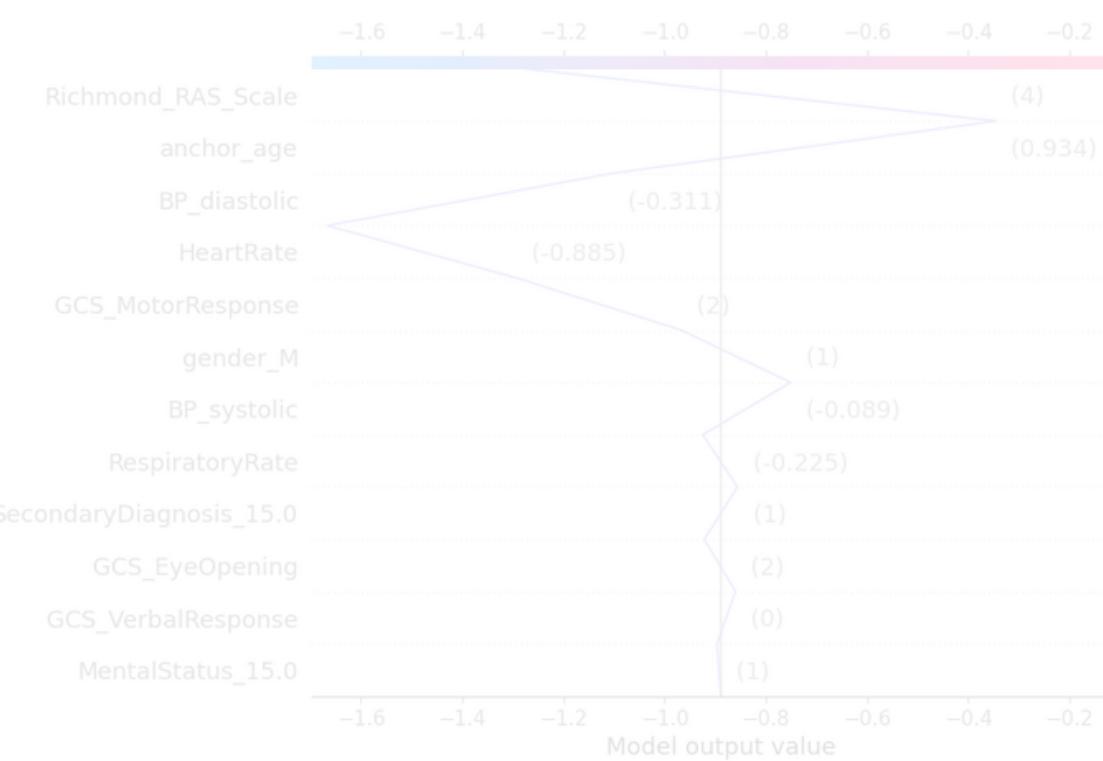
Modeling

Evaluation

Interpretation XAI-SHAP

Interpretation

SHAP (SHapley Additive exPlanation)



Data Preprocessing

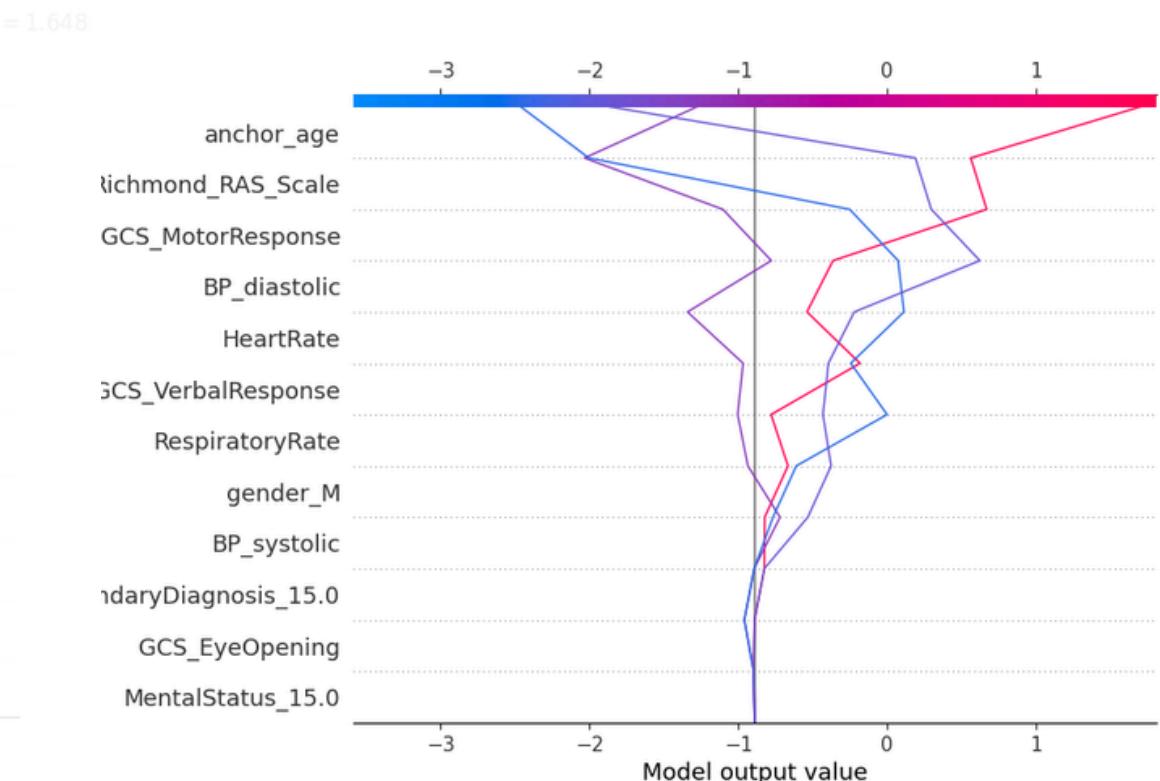
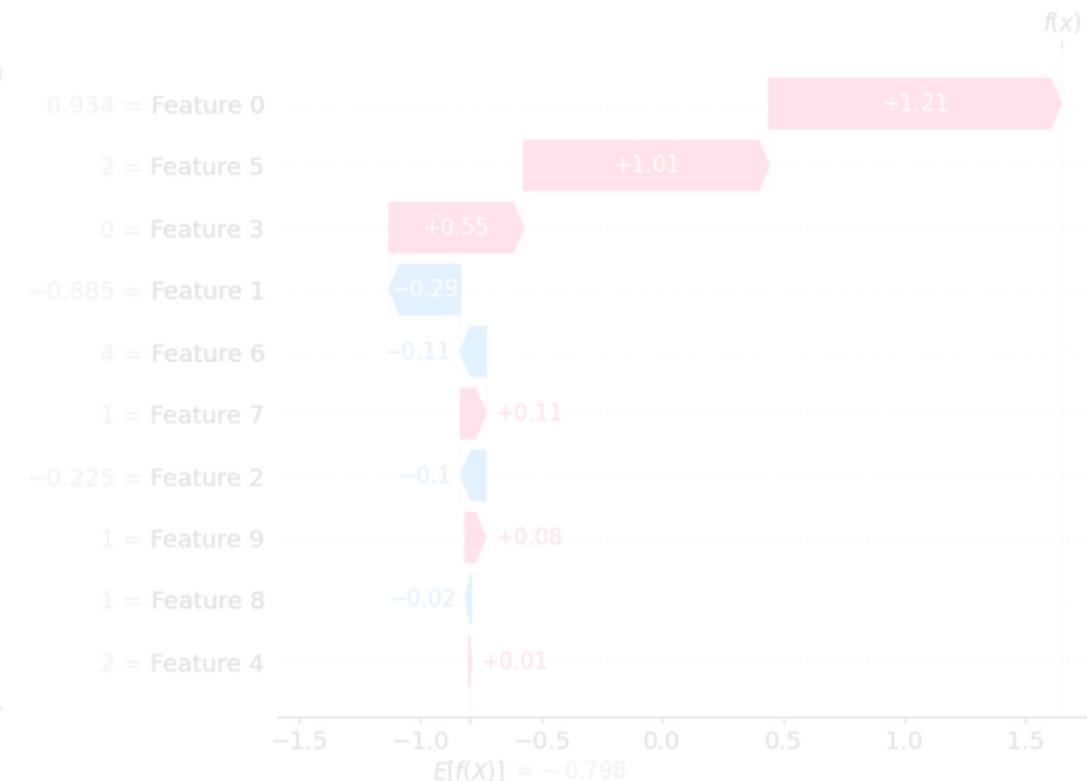
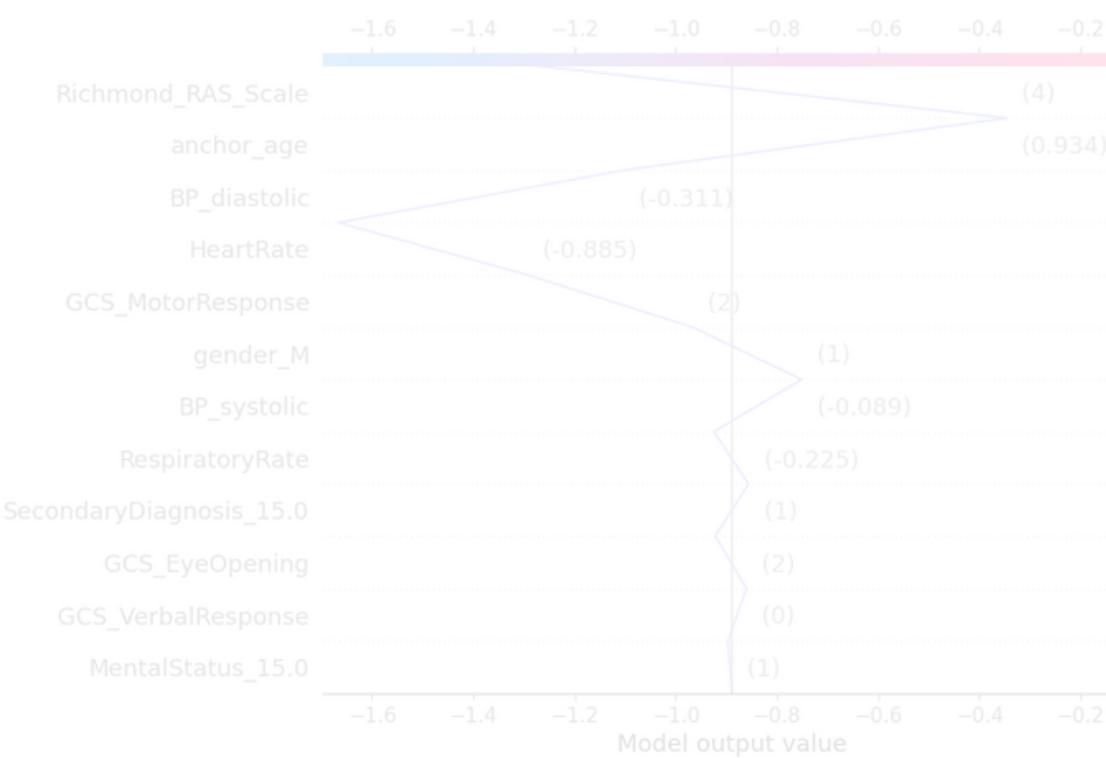
Modeling

Evaluation

Interpretation XAI-SHAP

Interpretation

SHAP (SHapley Additive exPlanation)



Data Preprocessing

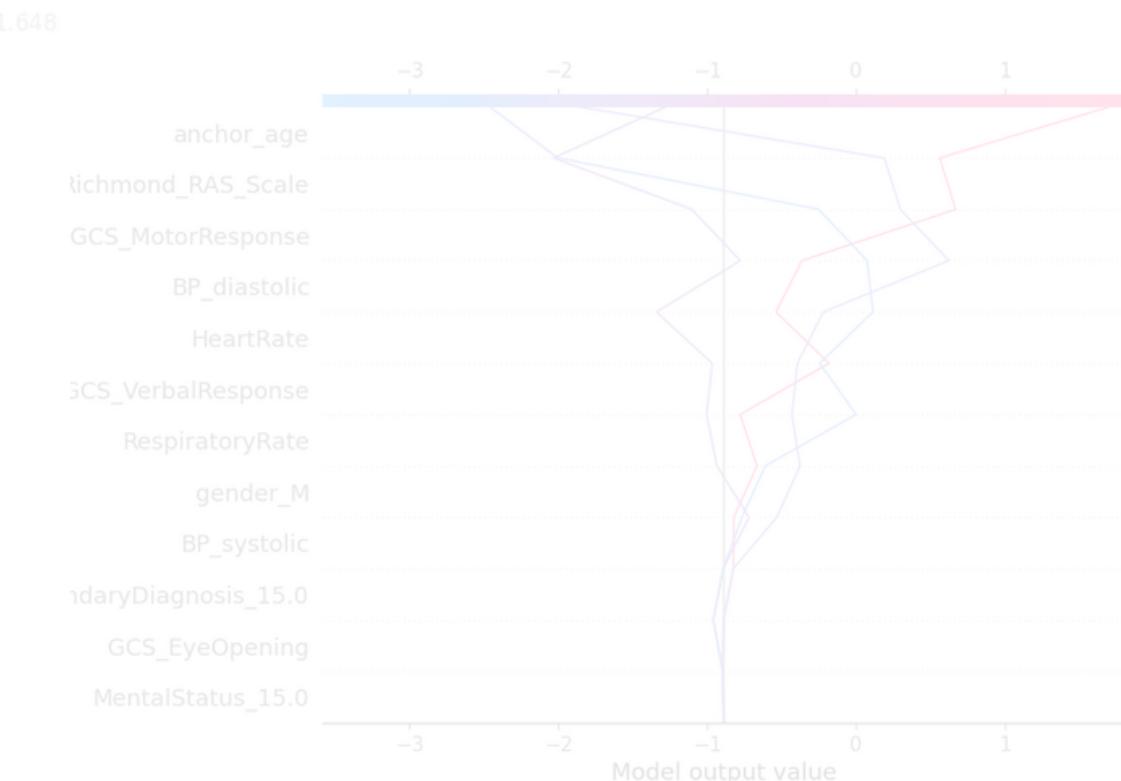
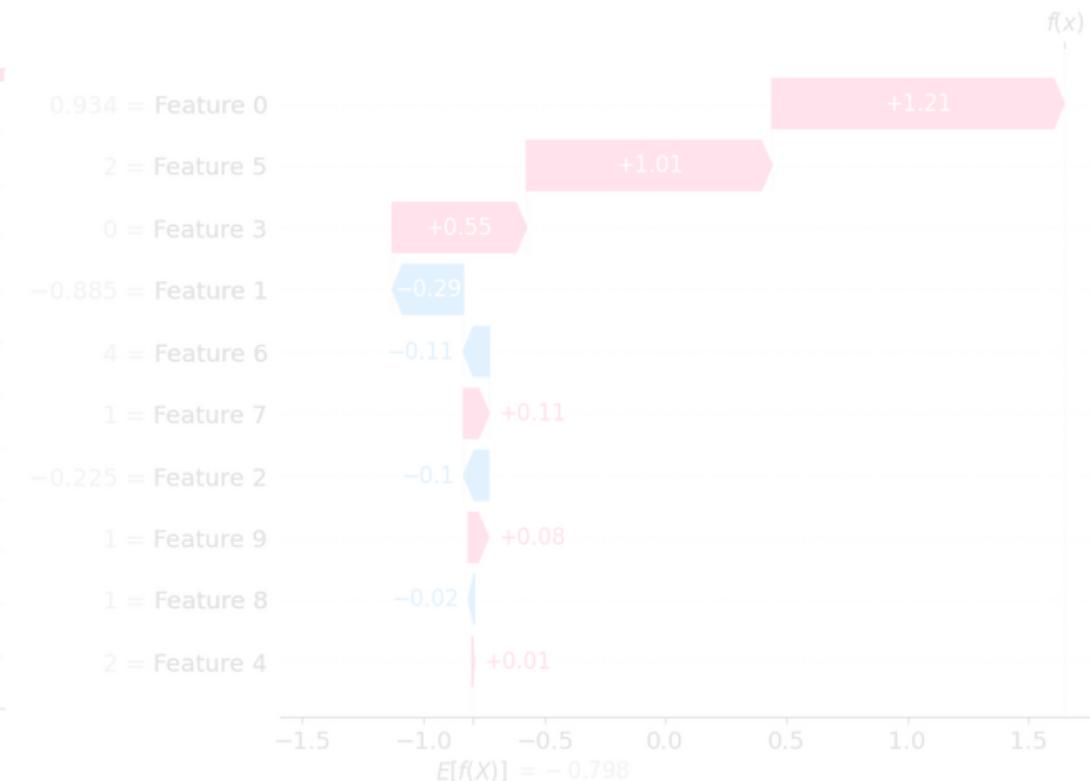
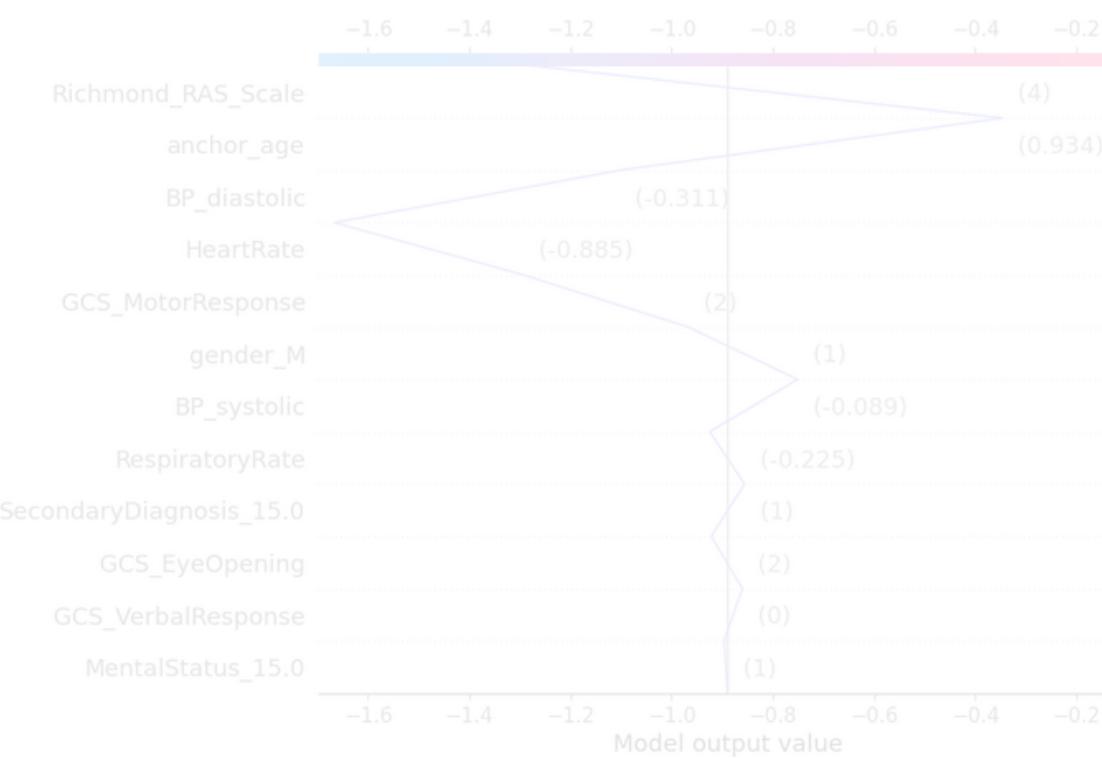
Modeling

Evaluation

Interpretation XAI-SHAP

Interpretation

SHAP (SHapley Additive exPlanation)



Data Preprocessing

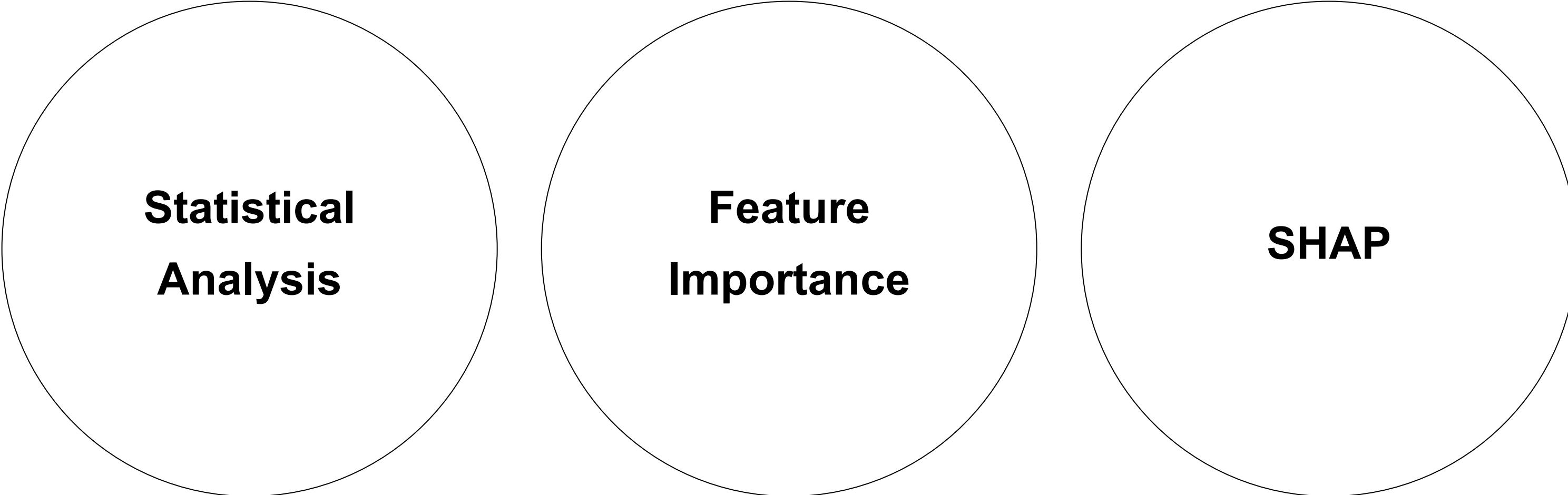
Modeling

Evaluation

Interpretation
XAI-SHAP

Interpretation

SHAP (SHapley Additive exPlanation)



**Statistical
Analysis**

**Feature
Importance**

SHAP

Interpretation

Nonlinear relationships

Complex interactions between data

Interpretation

Using multiple tools **organically together**

→ **Enhancing model performance and interpretability**

through **multi-faceted analysis**

Interpretation

Insignificant variables

→ When combined,

there is a potential for improved predictive power

Interpretation

Avoid creating derived variables unconditionally
→ Consider combinations **with and without** derived
variables

Data Preprocessing

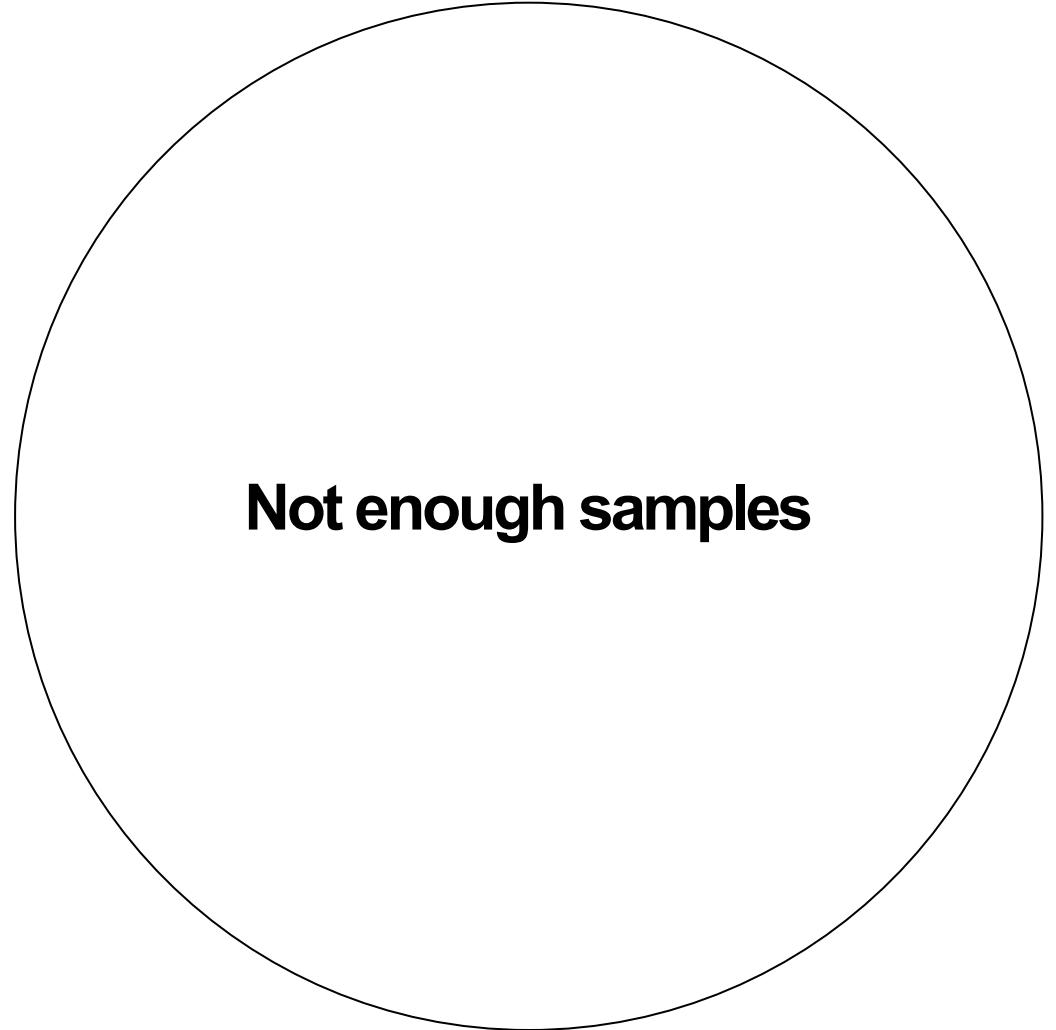
Modeling

Evaluation

Interpretation

Limitations / Solutions

Limitations / Solutions

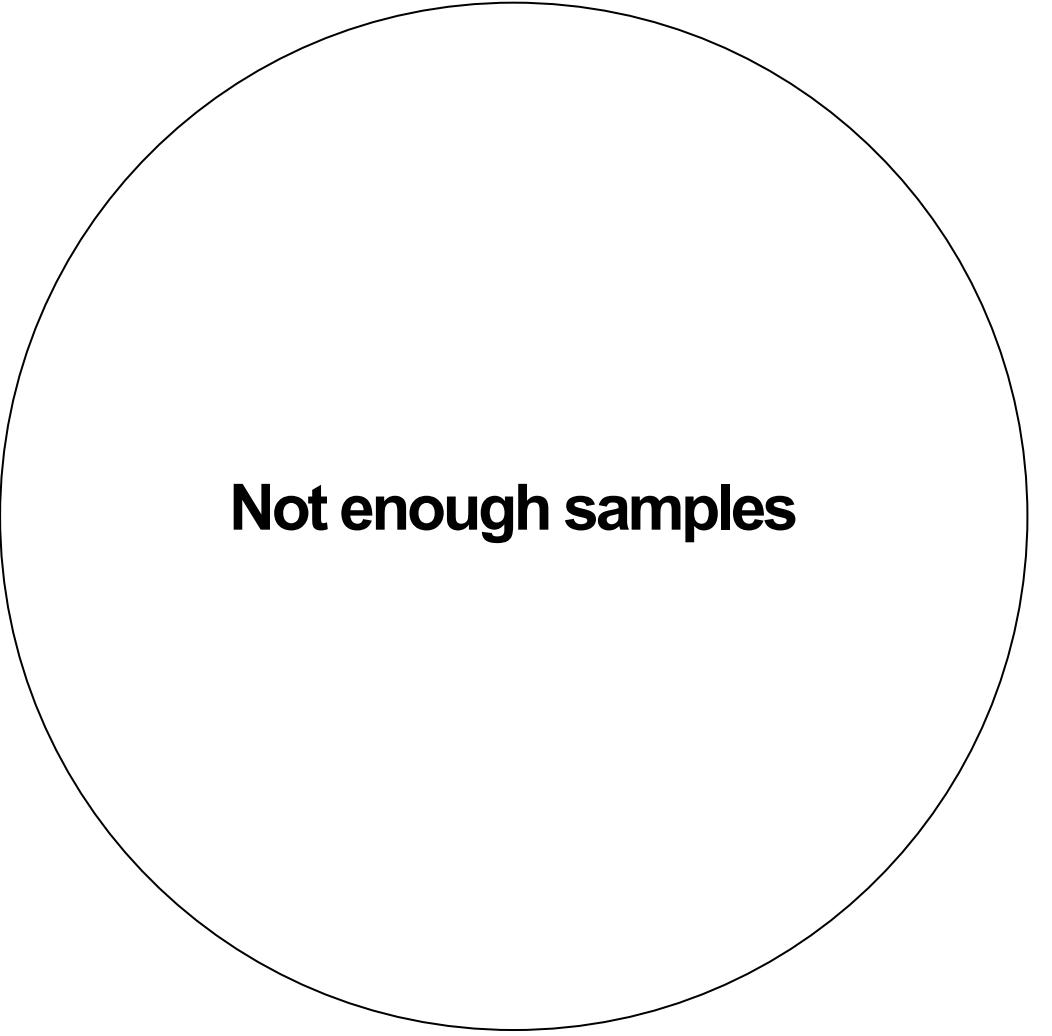


Not enough samples



Difficulty in collecting death samples

Limitations / Solutions

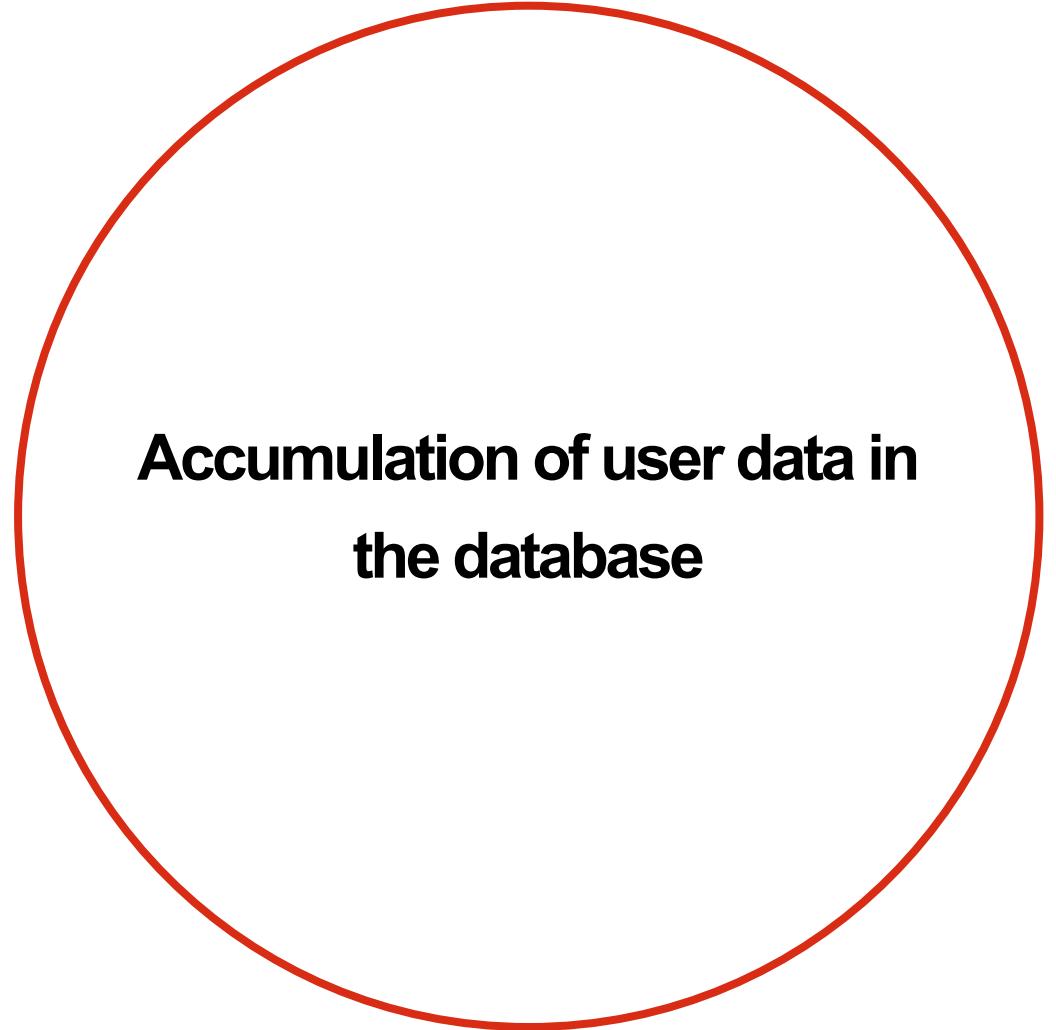


Not enough samples



사망 데이터 수집 어려움

Limitations / Solutions



**Accumulation of user data in
the database**



데이터 축적에 따라
새로운 모델 개발

Limitations / Solutions

데이터 수 부족

**Difficulty in collecting death
samples**

Data Preprocessing

Modeling

Evaluation

Interpretation

Limitations / Solutions

Limitations / Solutions

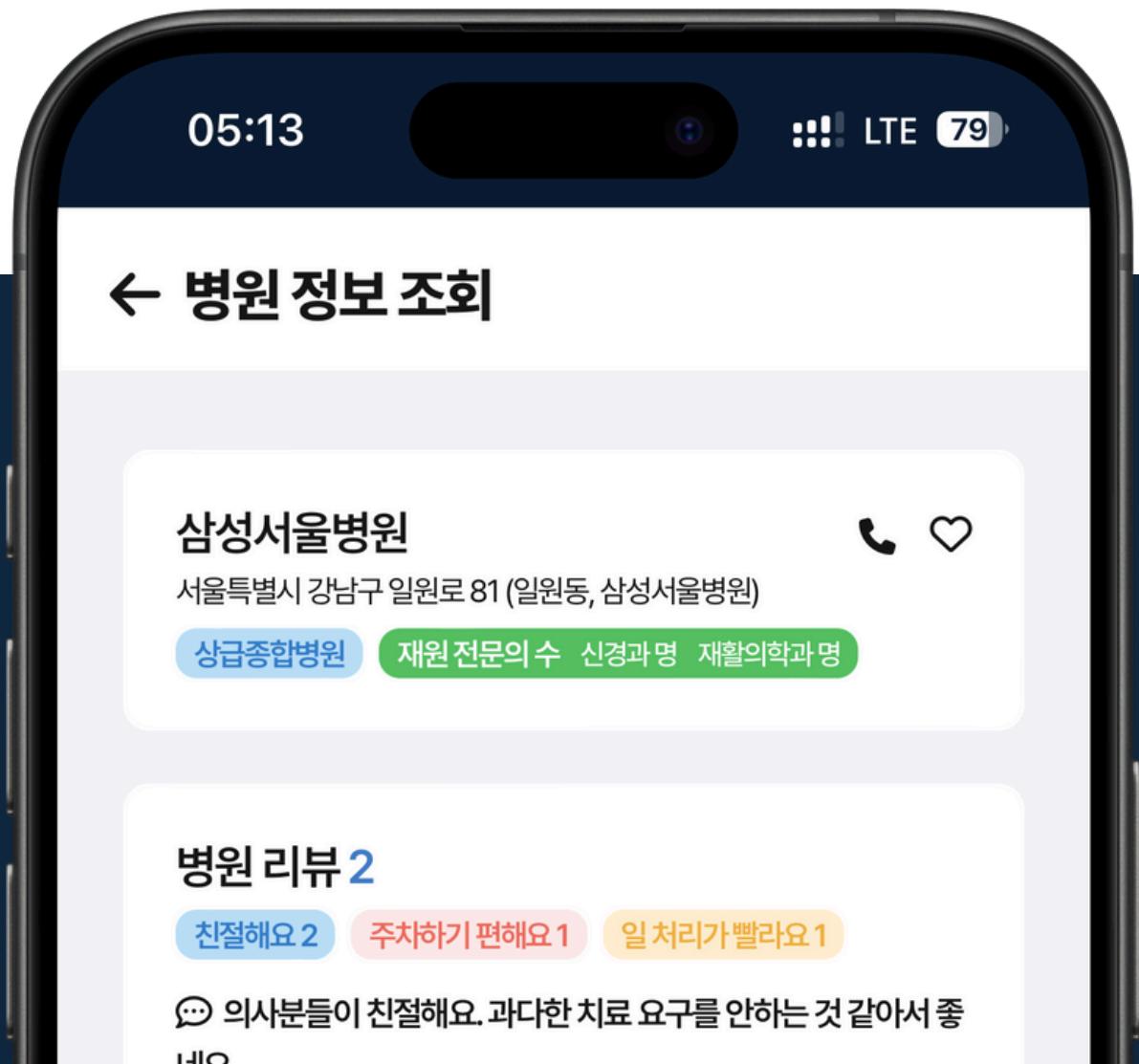
사용자 데이터 DB 축적

**Develop new models as data
accumulates**



Multi Platform Development

From responsive web/app implementation to final deployment



Development Dept

Jigeon Park (Full-stack dev)

Overview

The development team's goal is to **enhance user experience** by optimizing the **stability and performance** of the product.

Overview

Agile Methodology

: A flexible approach that involves releasing products quickly in regular cycles, allowing for the addition and modification of requirements to meet customer needs and adapt to changing environments.

Dev Lang. / Frameworks



python

django

Database / Manage



MariaDB

phpMyAdmin

CI/CD



git



GitHub

Web Server, Proxy

NGINX

Dev environment, IDE



mac
OS



PyCharm



VSCode

ML Pipeline



JobLib

Server Protection, DNS



Route



NGINX
PROXY MANAGER

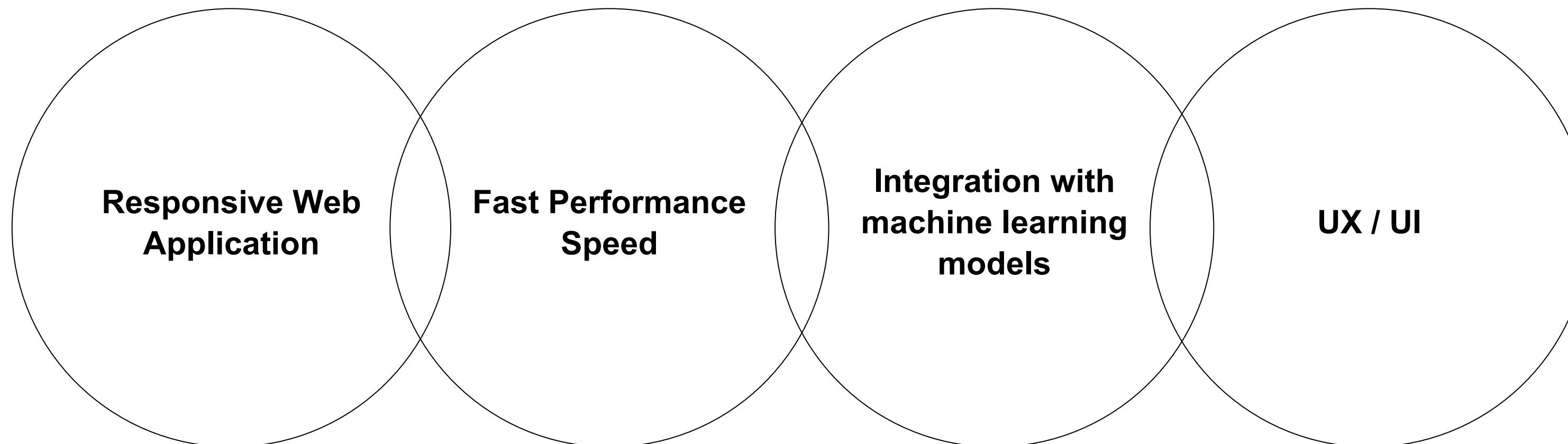
Server



Container Orchestration



Goals



Tasks

Integration with multiple APIs

Speed and maintenance of machine learning models

Library bugs

Tasks

Integration with multiple APIs

머신러닝 모델의 속도 및 유지보수

라이브러리 버그

Tasks

Integration with multiple APIs

**When querying the hospital information list and detailed information,
30 to 180 API calls occur**

Tasks

Integration with multiple APIs

Causing an **unusable level of speed reduction**

*Each query requires at least 130 seconds

Tasks

Integration with multiple APIs

머신
속도

**After applying caching
technology,
speed improves by at least 130 to 1,200 times**

Tasks

**Speed and maintenance
of ML models**

다중 API 연동문제

라이브러리 버그

Tasks

Speed and maintenance
of ML models

**Operating and integrating ML models on a
service server with lower performance
compared to GCP**

Tasks

Speed and maintenance
of ML models

Continuously maintaining the **improved**
다중 API 연동문제
version of the new ML models
라이브러리 버그

Tasks

**Speed and maintenance
of ML models**

**Using the *Joblib* library,
to integrate the built ML model pipeline, solving
both speed and maintenance issues**

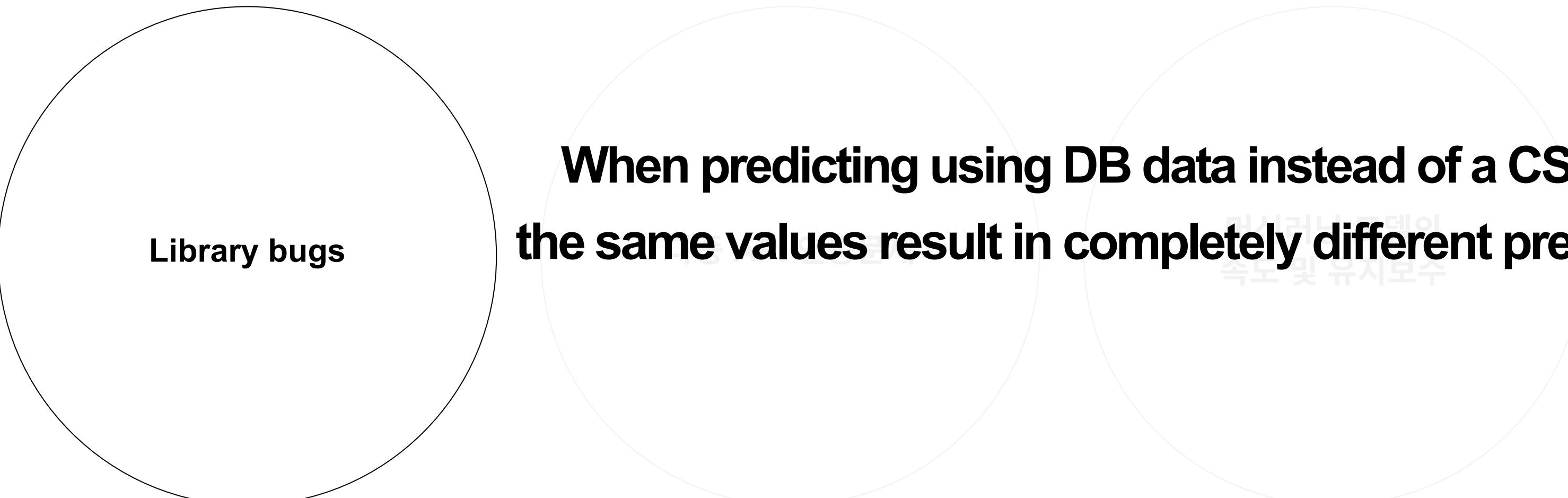
Tasks

다중 API 연동문제

머신러닝 모델의
속도 및 유지보수

Library bugs

Tasks



Library bugs

**When predicting using DB data instead of a CSV file,
the same values result in completely different predictions**

목표 및 사용자 수

Tasks

Library bugs

Export to CSV, then reload it

Manage the CSV files using unique user codes, and
configure the system to delete the CSV file
immediately after loading is complete

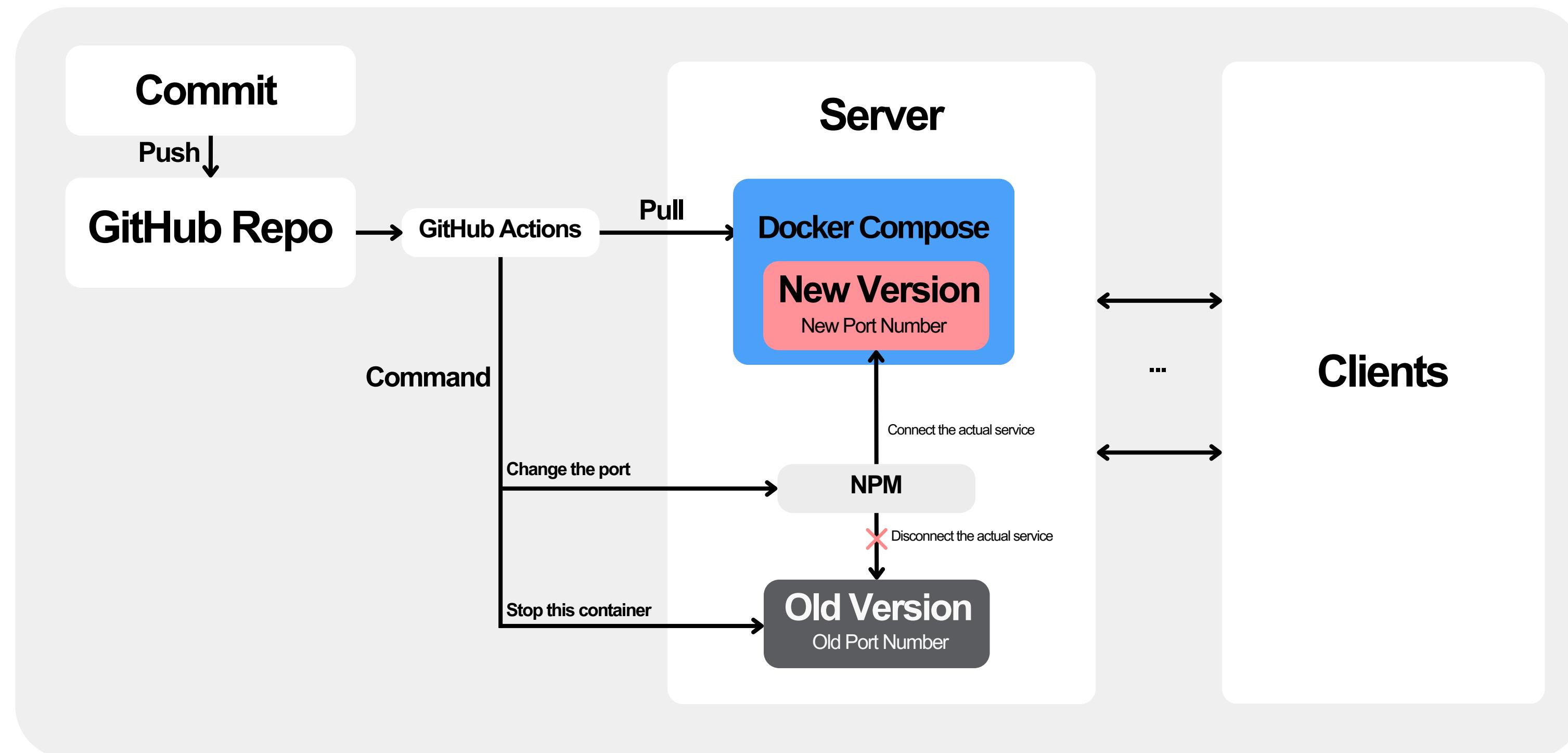
Considerations

Handling Frequent Changes

Full automation of the entire process

Uninterrupted deployment and quick speed

Flow chart



Flow chart



Flow chart

Commit 발생
Push ↓
GitHub Repo → Docker Compose
Pull → New Version → New Port Number
Change the port → Connect the actual service → Disconnect the actual service
Stop this container → Old Version
Old Port Number

**Compared to before automation,
shows a minimum **35-fold** improvement in
efficiency**

Overview

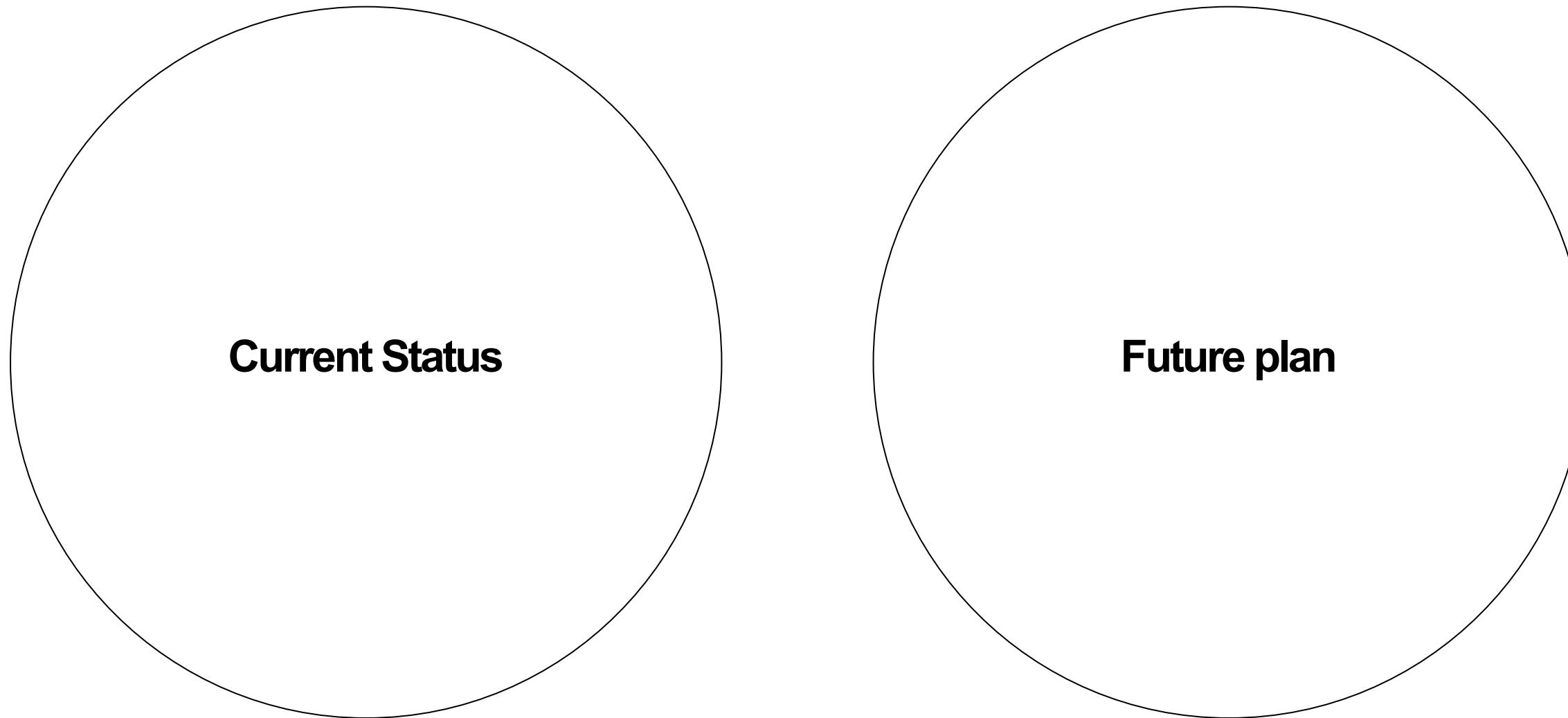
Technologies

Goals / Tasks

Operational plan

Maintenance plan

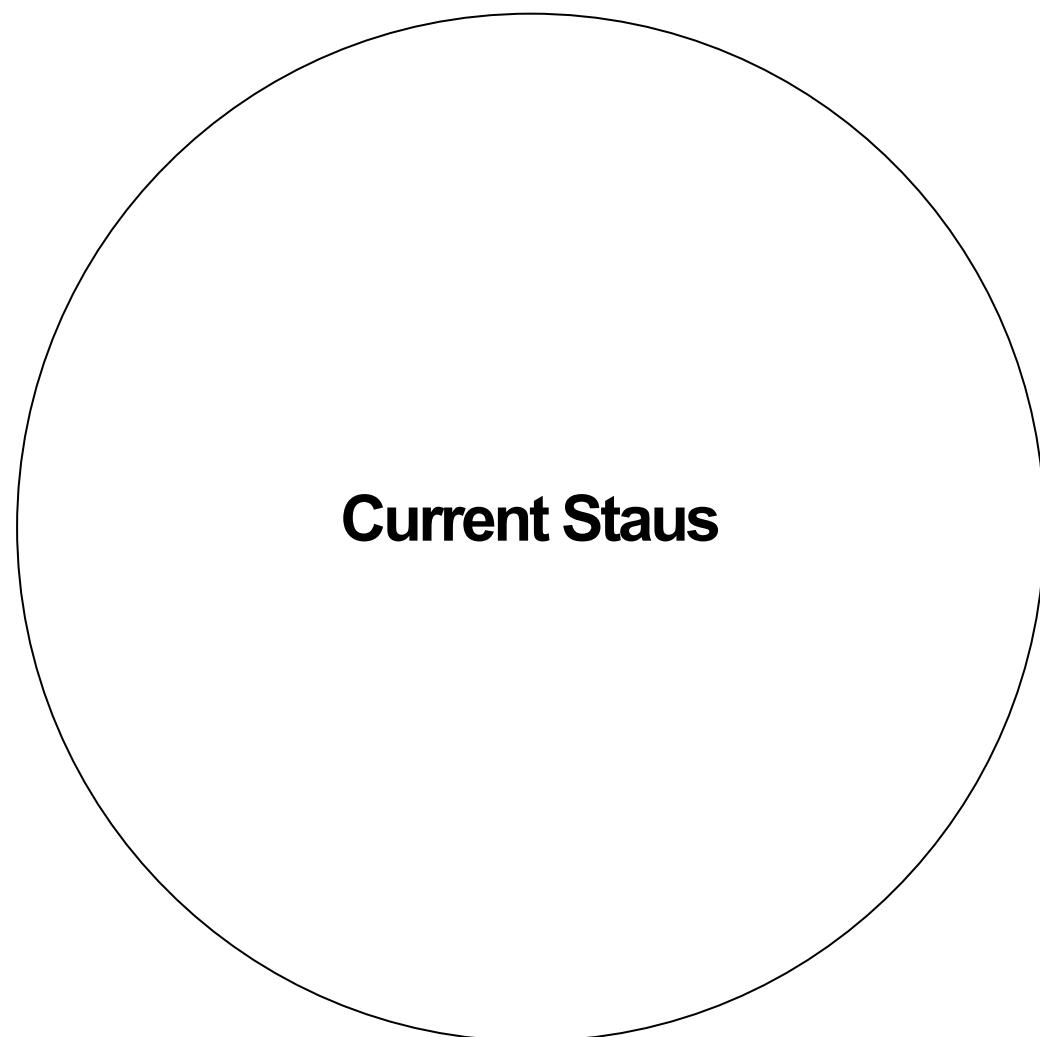
Maintenance plan



Current Status

Future plan

Maintenance plan



Current Status

Integration with Google Analytics

Track the volume of incoming traffic/simultaneous users
Analyse user trends and behaviors

Overview

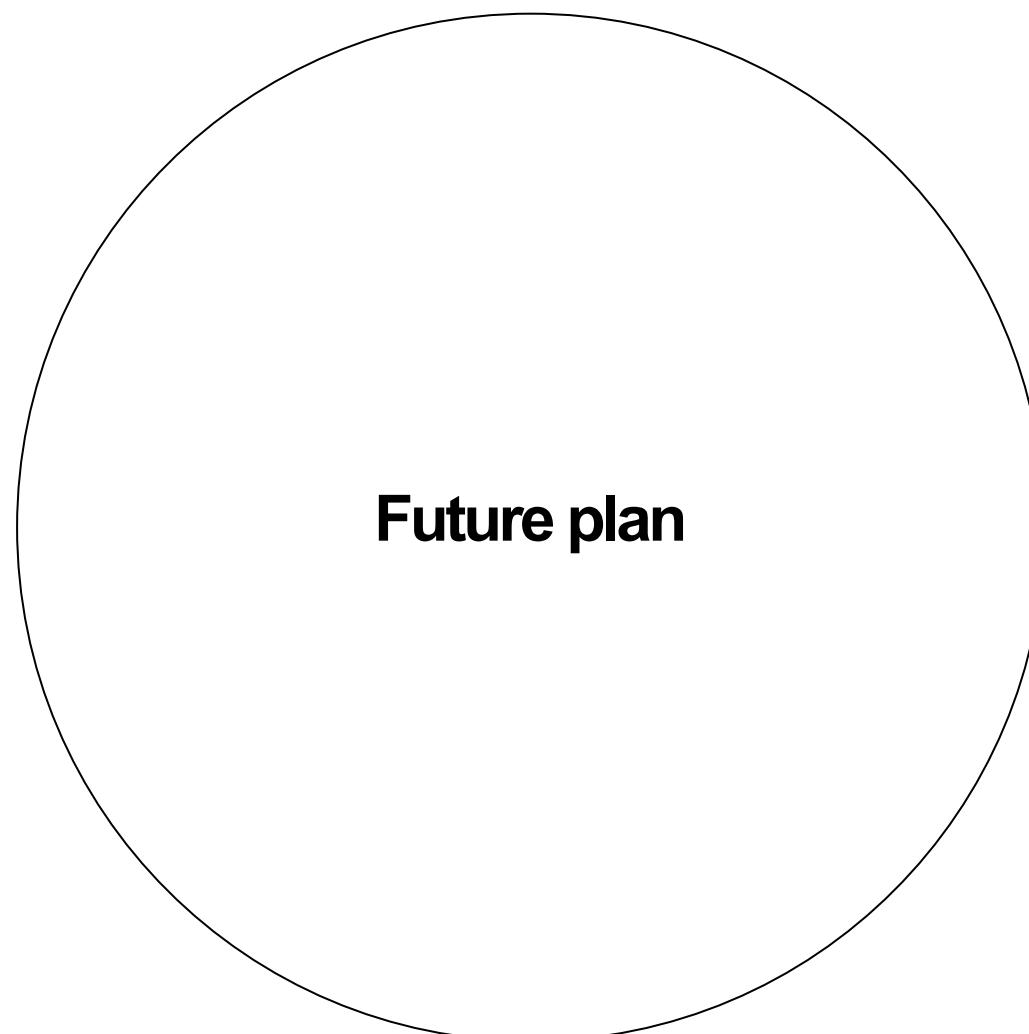
Technologies

Goals / Tasks

Operational plan

Maintenance plan

Maintenance plan



 Prometheus  Grafana

**Utilize Advanced Monitoring
tool**

For advanced user behavior analysis

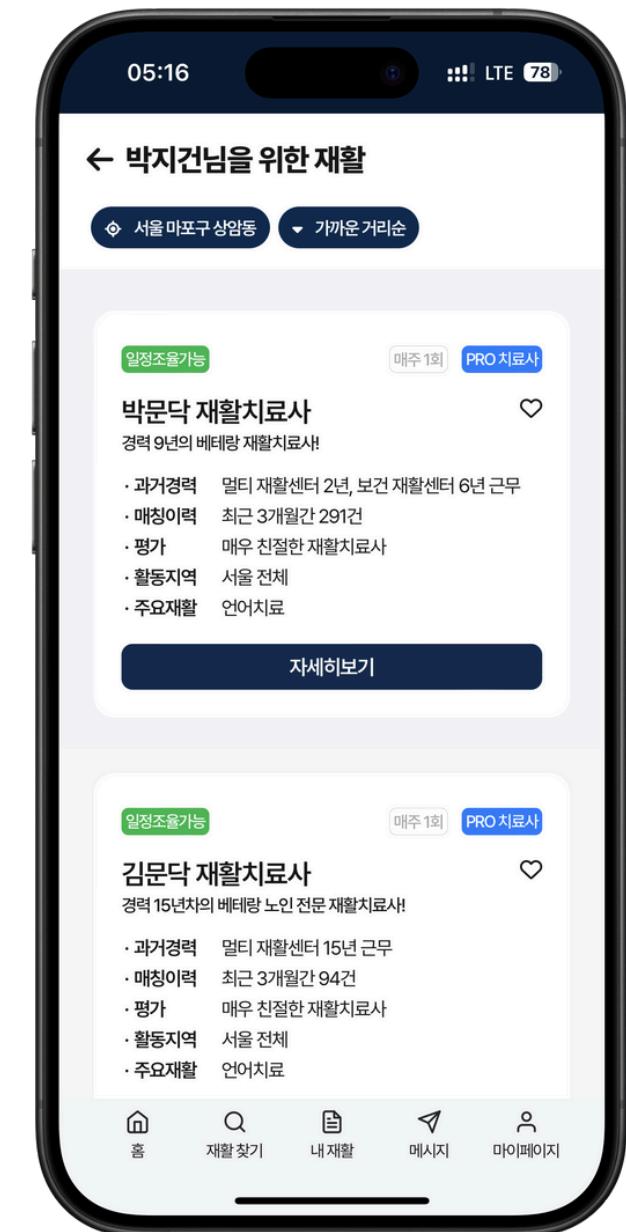
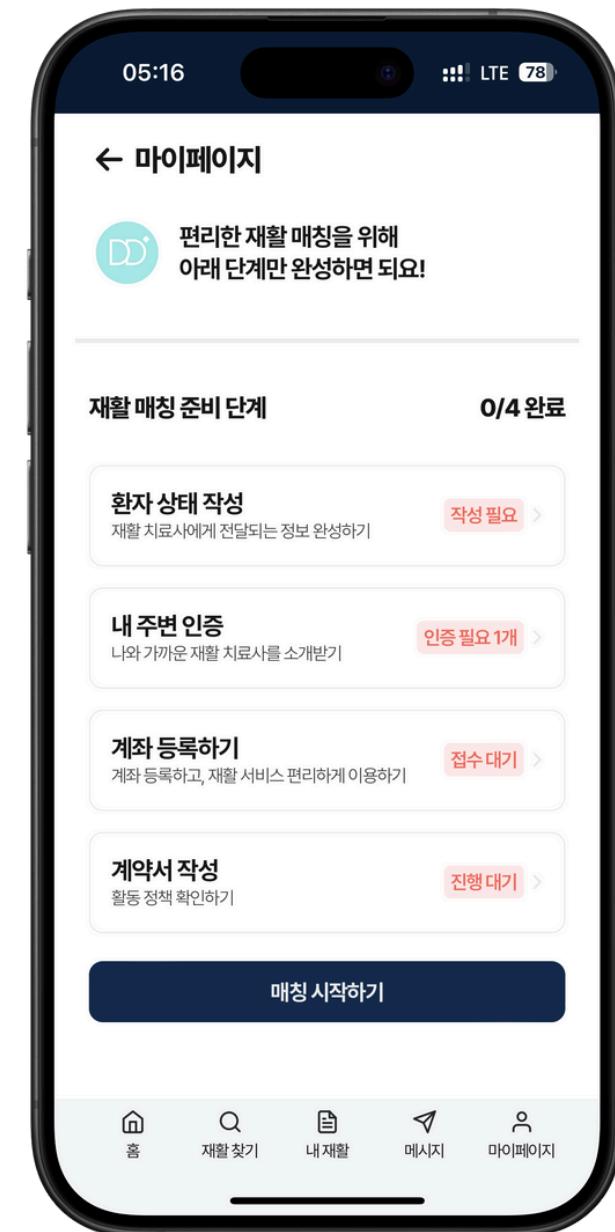
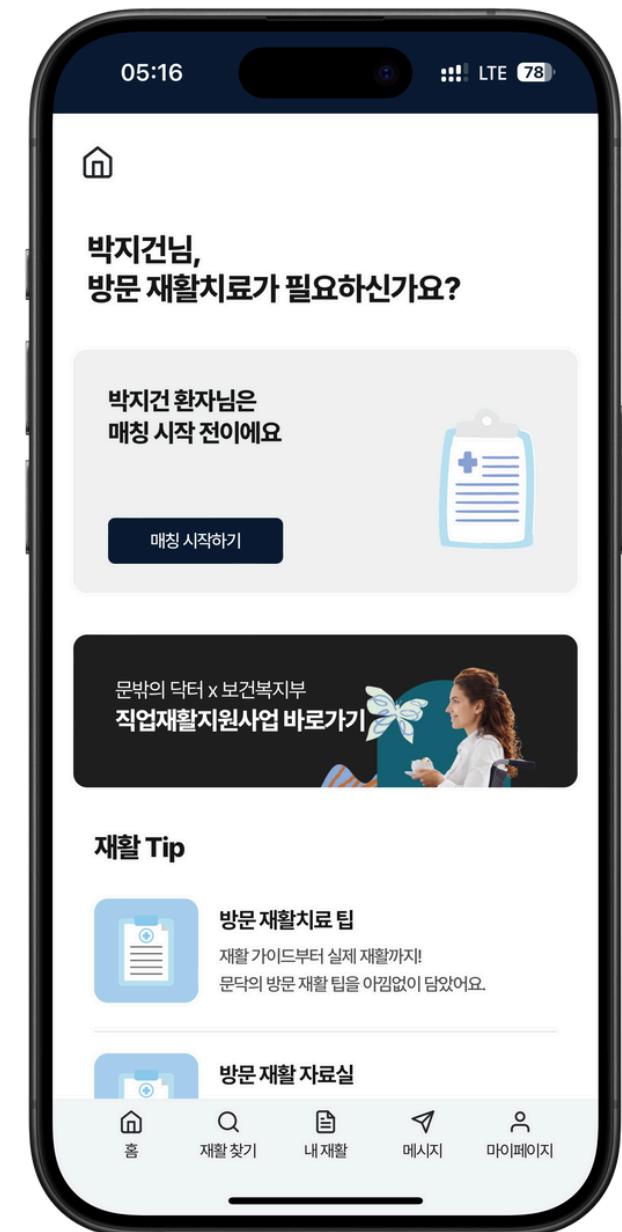
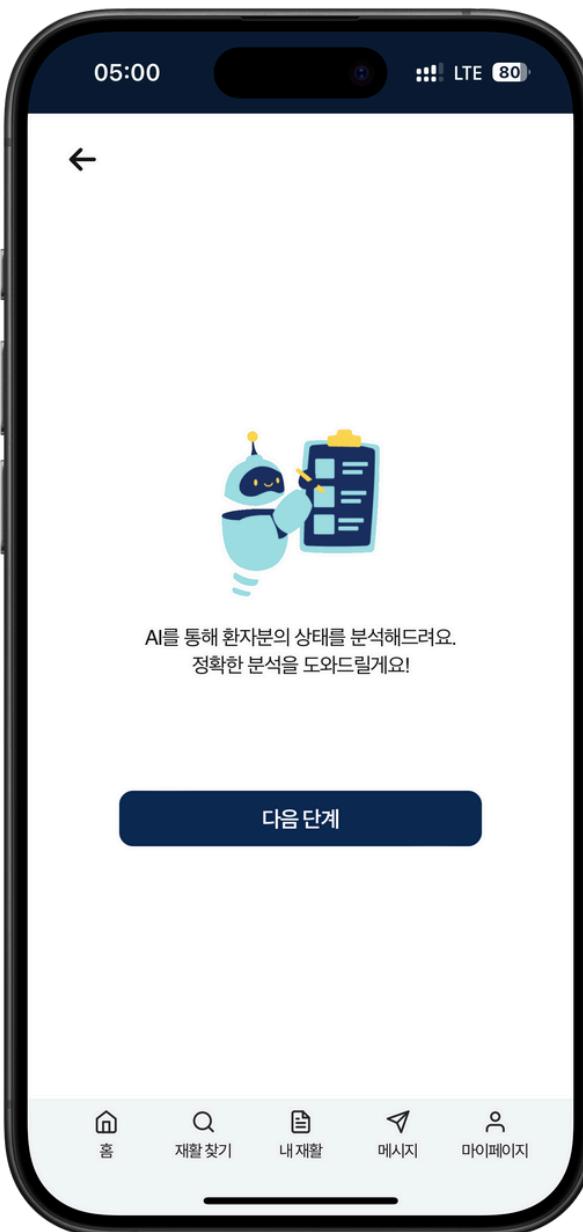
Maintenance plan



추후 계획

Migration to Cloud service
Enhance service stability through cloud migration
to AWS, Google Cloud Platform

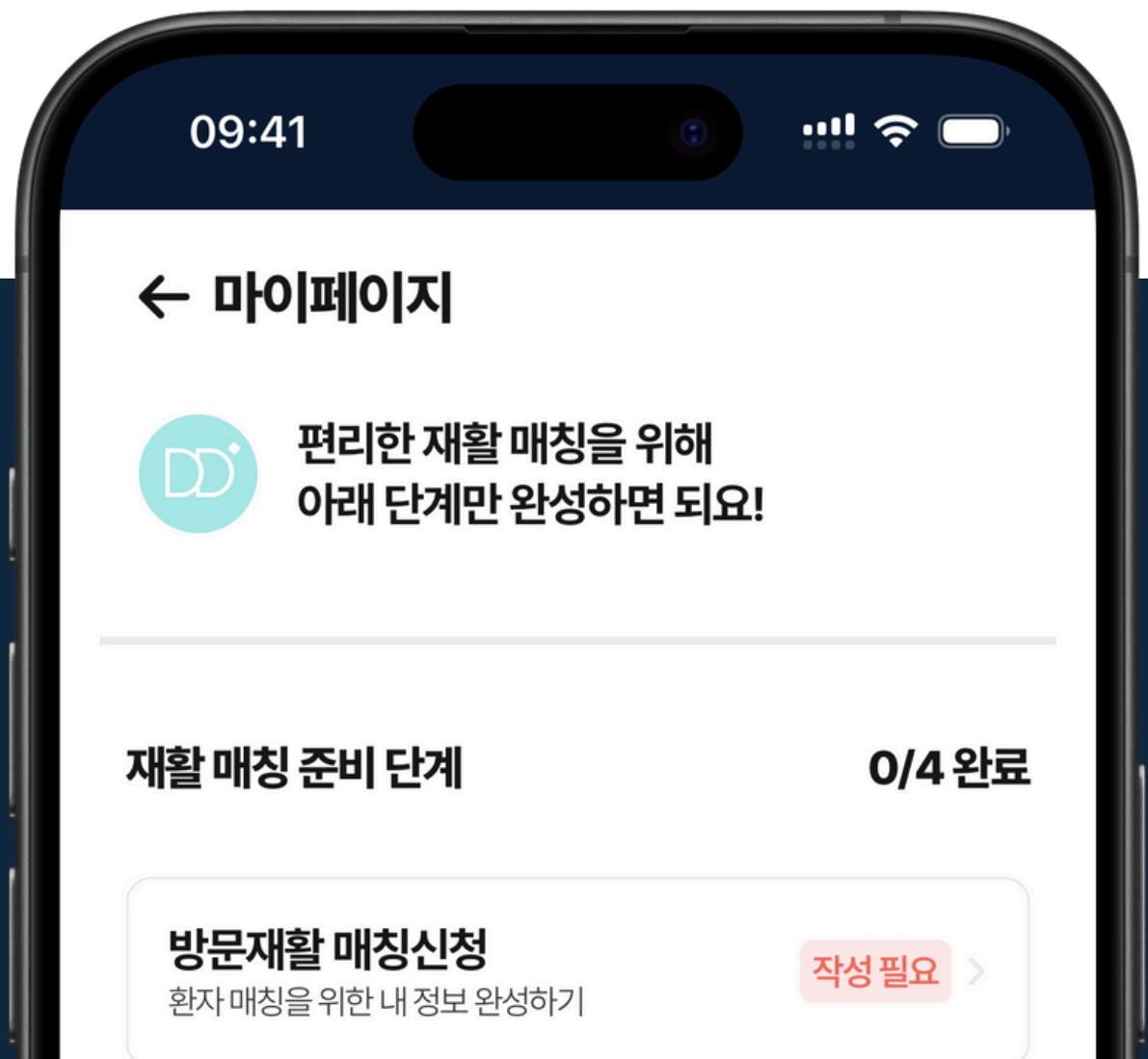
Multi Platform Demo





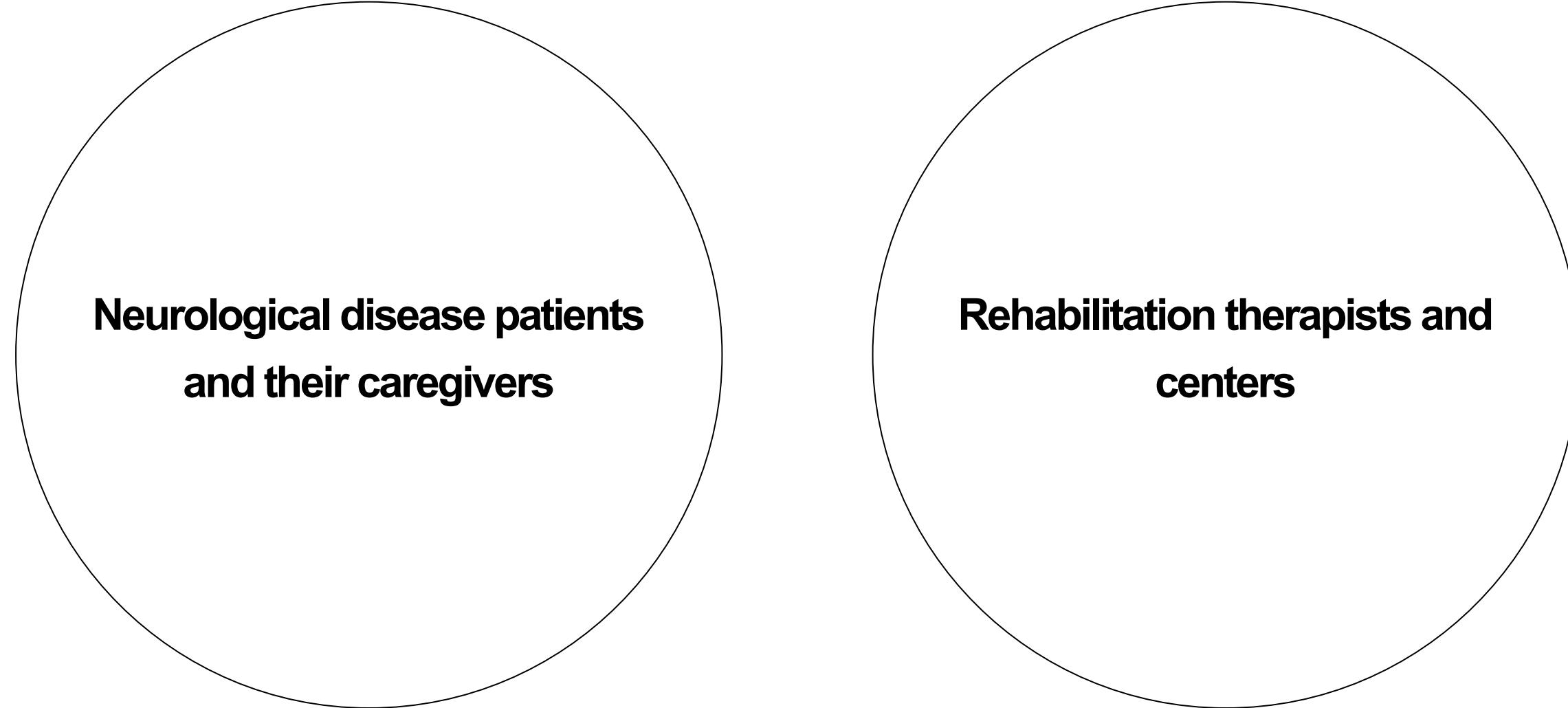
Business Model

Market & Target Analysis / Expected Outcomes



Business Planning Dept
Hyewon Jung
Jihoon Youn

Market/Target Analysis



**Neurological disease patients
and their caregivers**

**Rehabilitation therapists and
centers**

Market/Target Analysis



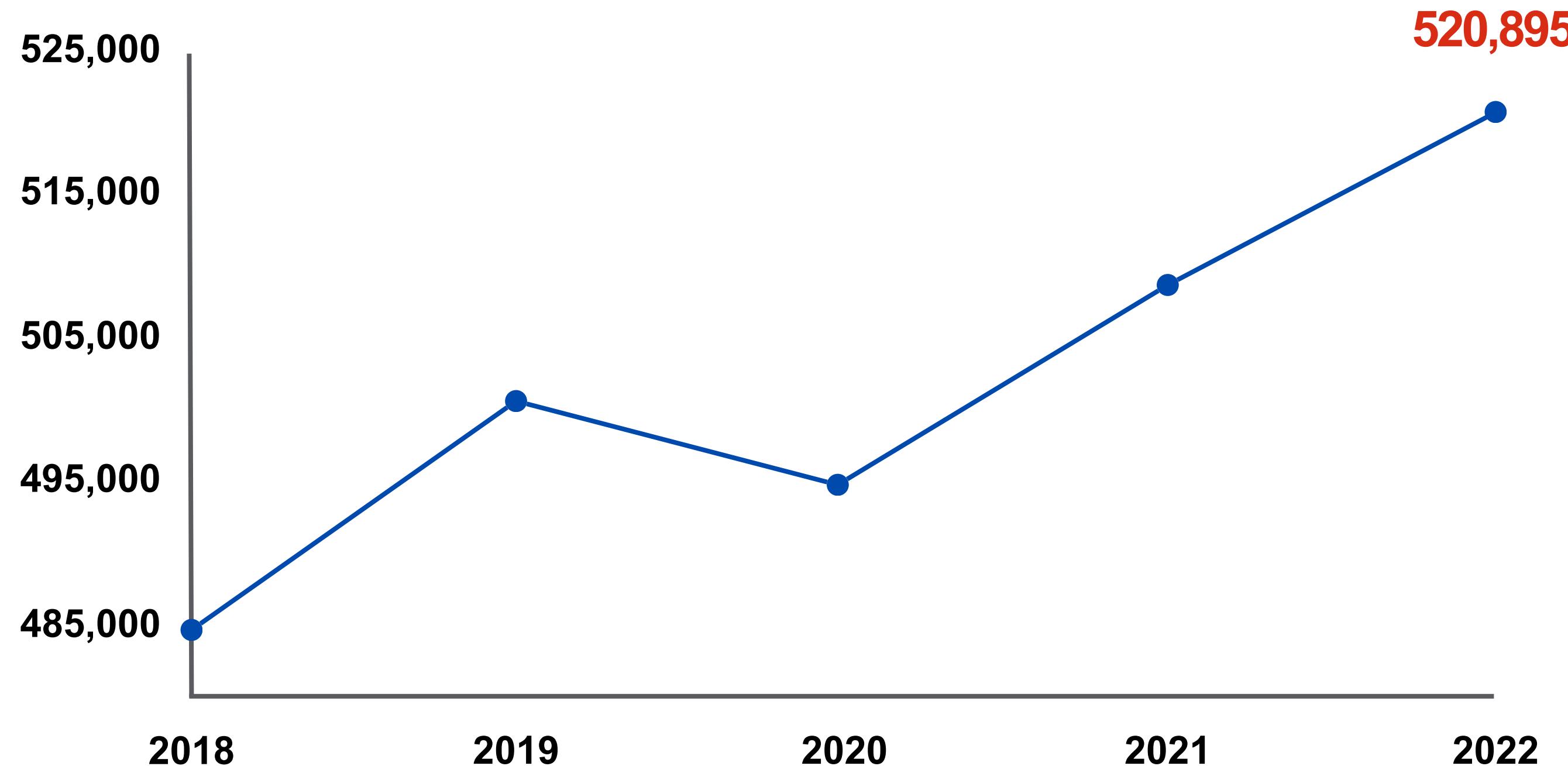
**Neurological disease patients
and their caregivers**

Population size : Approx. 128 million

Market size : 5.26 billion dollars

Market/Target Analysis

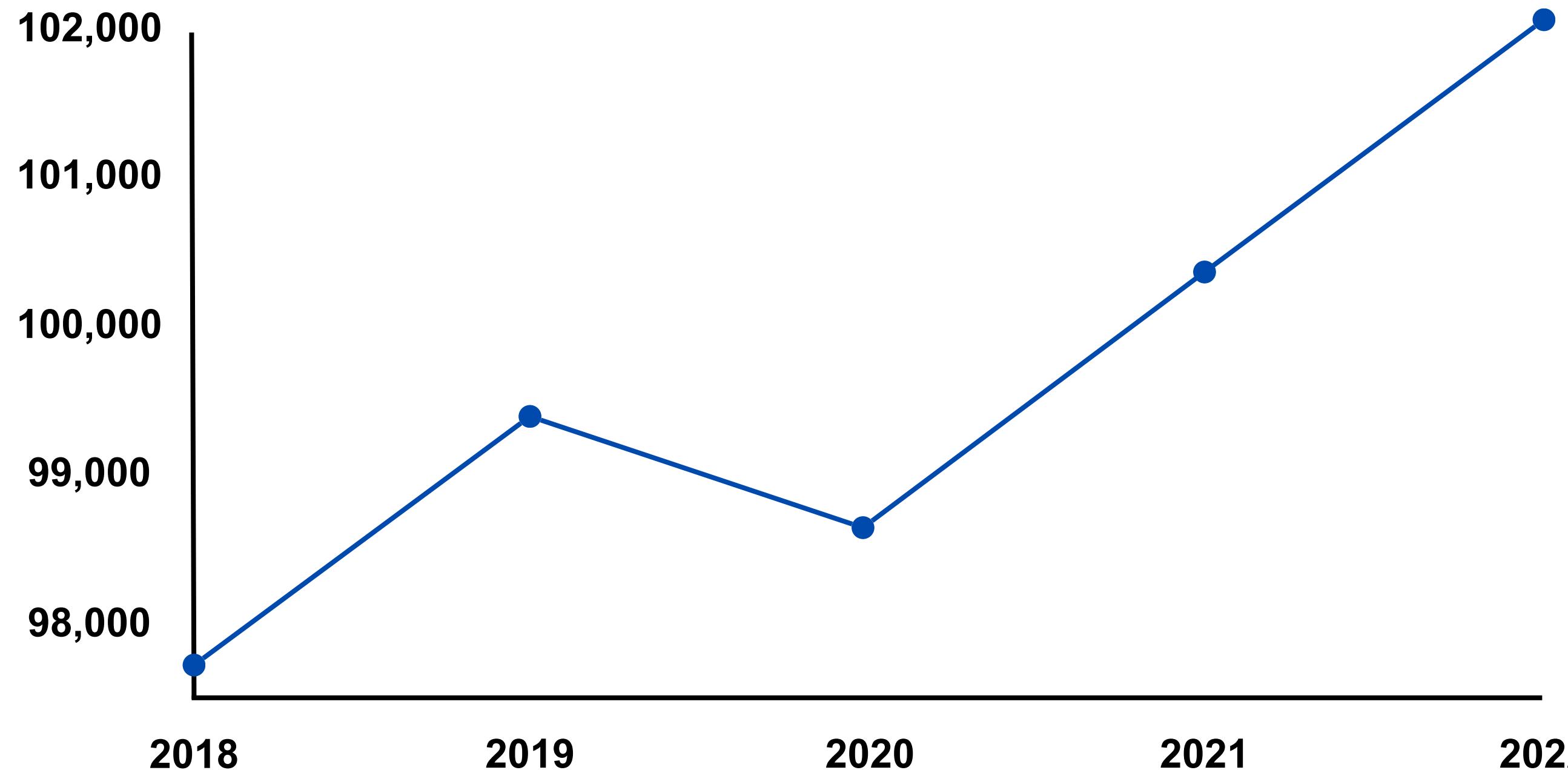
Statistics of patients with Cerebral infarction



Market/Target Analysis

Statistics of patients with Cerebral hemorrhage

102,127

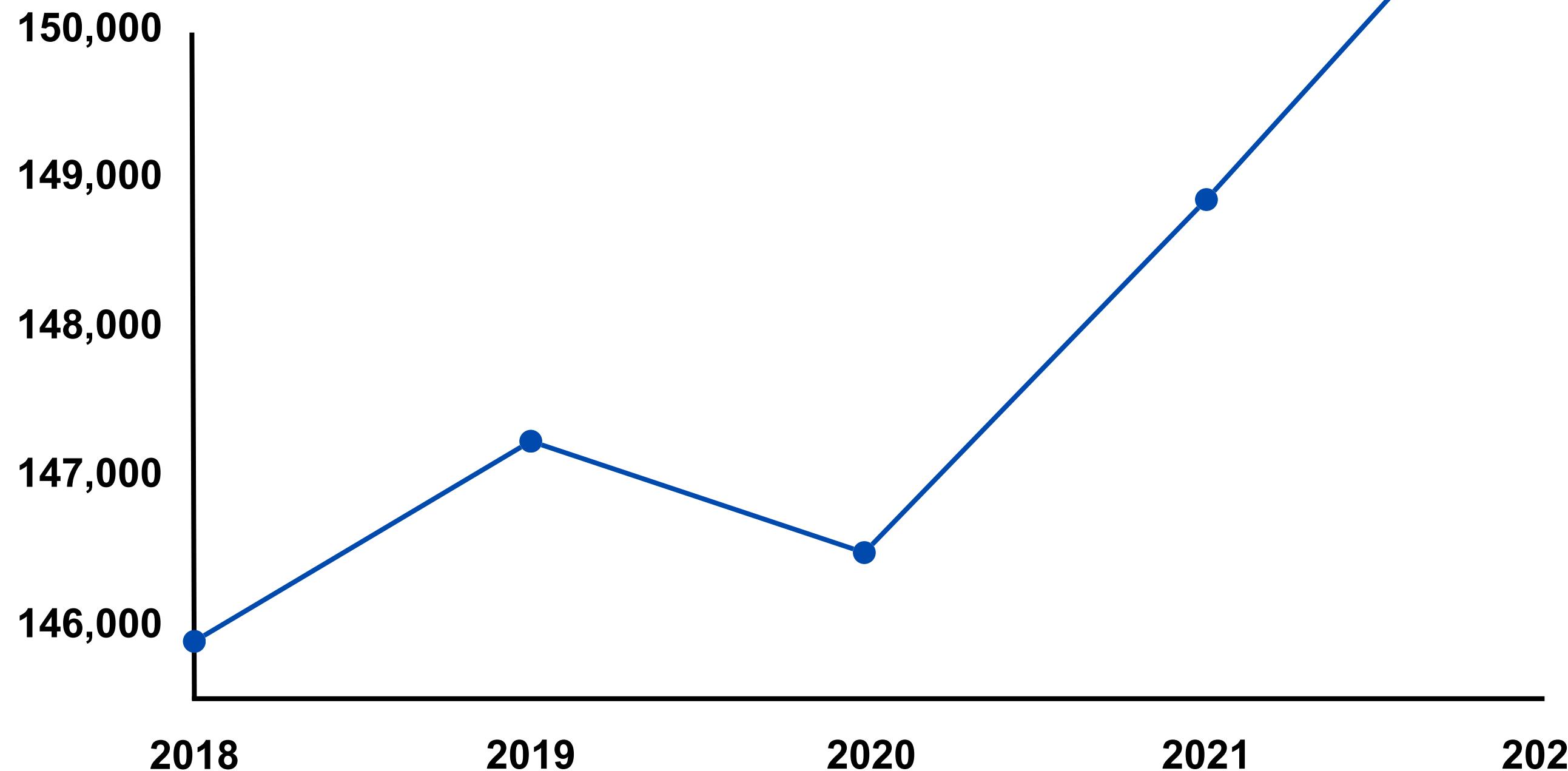


Market/Target Analysis

Expected Outcomes

Market/Target Analysis

Statistics of patients with Epilepsy



Market/Target Analysis

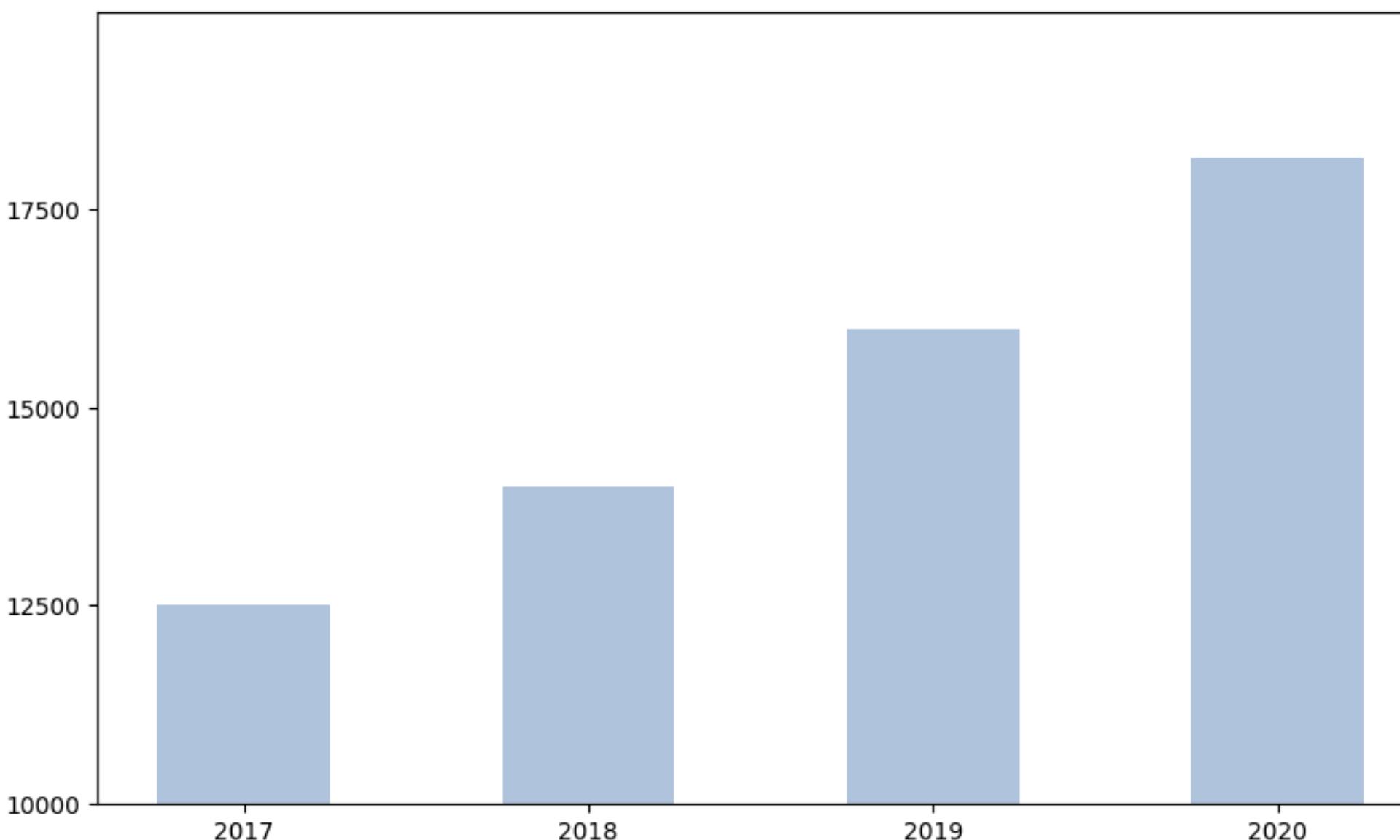


**Population size : approx. 21,000,
430 centers**

**Market size : approx. 1.56 billion
dollars**

Market/Target Analysis

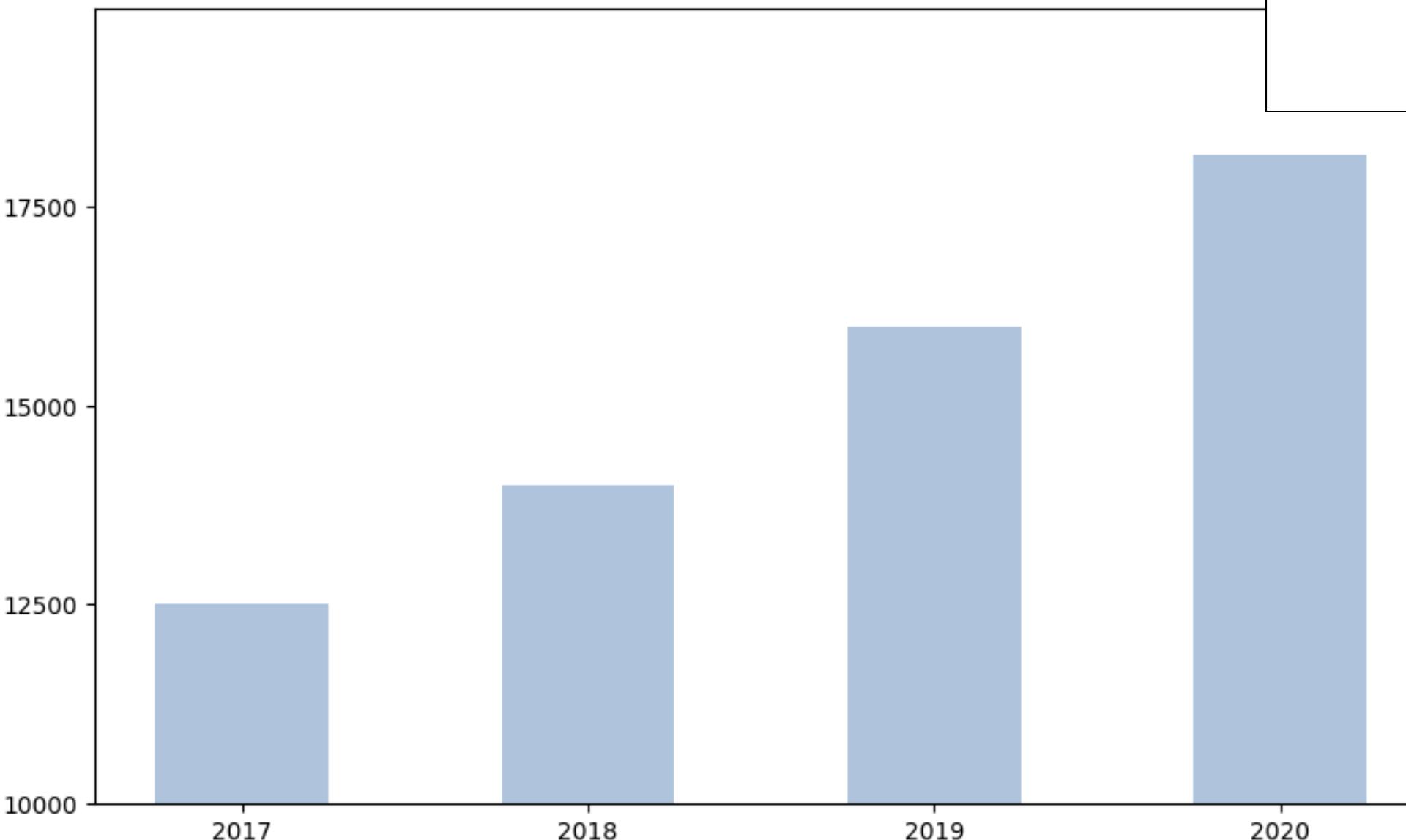
Numbers of Occupational therapist



Market/Target Analysis

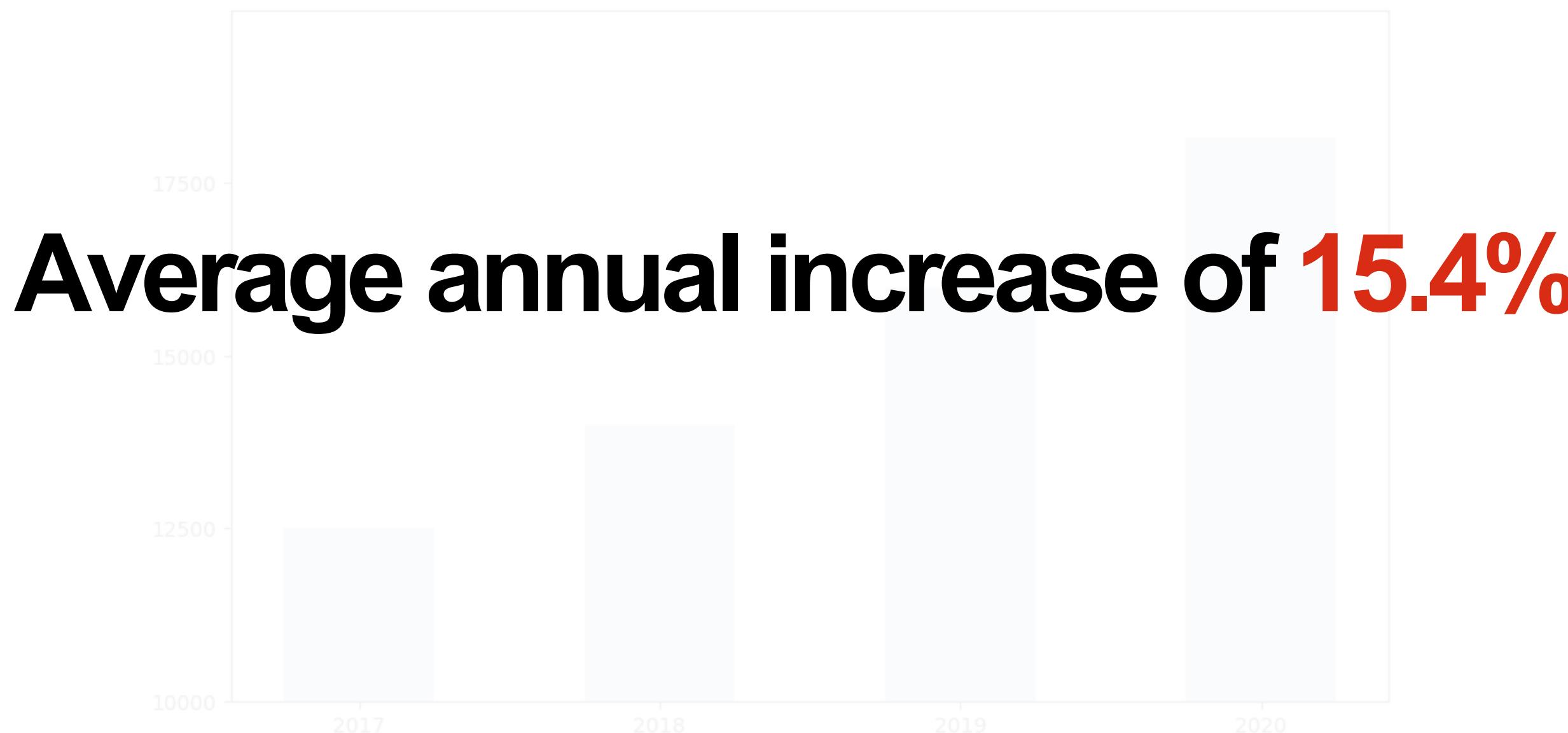
Numbers of Occupational therapist

Over 18,000 in year 2020

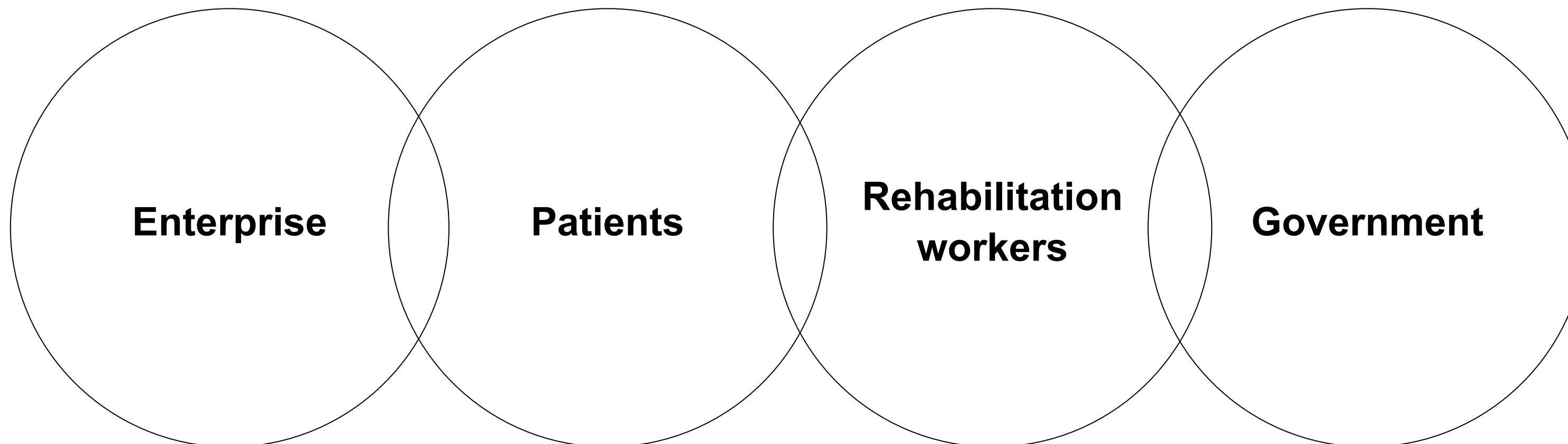


Market/Target Analysis

Numbers of Occupational therapist



Expected Outcomes



Expected Outcomes

Enterprise

By leveraging accumulated data, enterprises
can improve the **accuracy of AI models** and
enhance the quality of analytical results

Expected Outcomes

Enterprise

Providing a developed welfare infrastructure
can improve the **quality of life** for citizens

Expected Outcomes

Patients

Enables quick access to necessary
information and the design of optimal
rehabilitation plans

Expected Outcomes

Patients

Maximizes the effectiveness of rehabilitation
and reduces **medical expenses**

Expected Outcomes

Rehabilitation
workers

Flexible work is possible with segmented time
blocks, additionally generating extra revenue

Expected Outcomes

Government

Creates **21,000 new rehabilitation therapy jobs**
and stimulates economic activity amounting
to **1.12 billion dollars annually**

Expected Outcomes

Government

Achieves **economic stimulation** amounting to
1.12 billion dollars annually

References

<thesis>

MINYOUNG LEE. (2022). Machine learning based models for readmission to intensive care units. Graduate School of Management of Technology Korea University

<statistics>

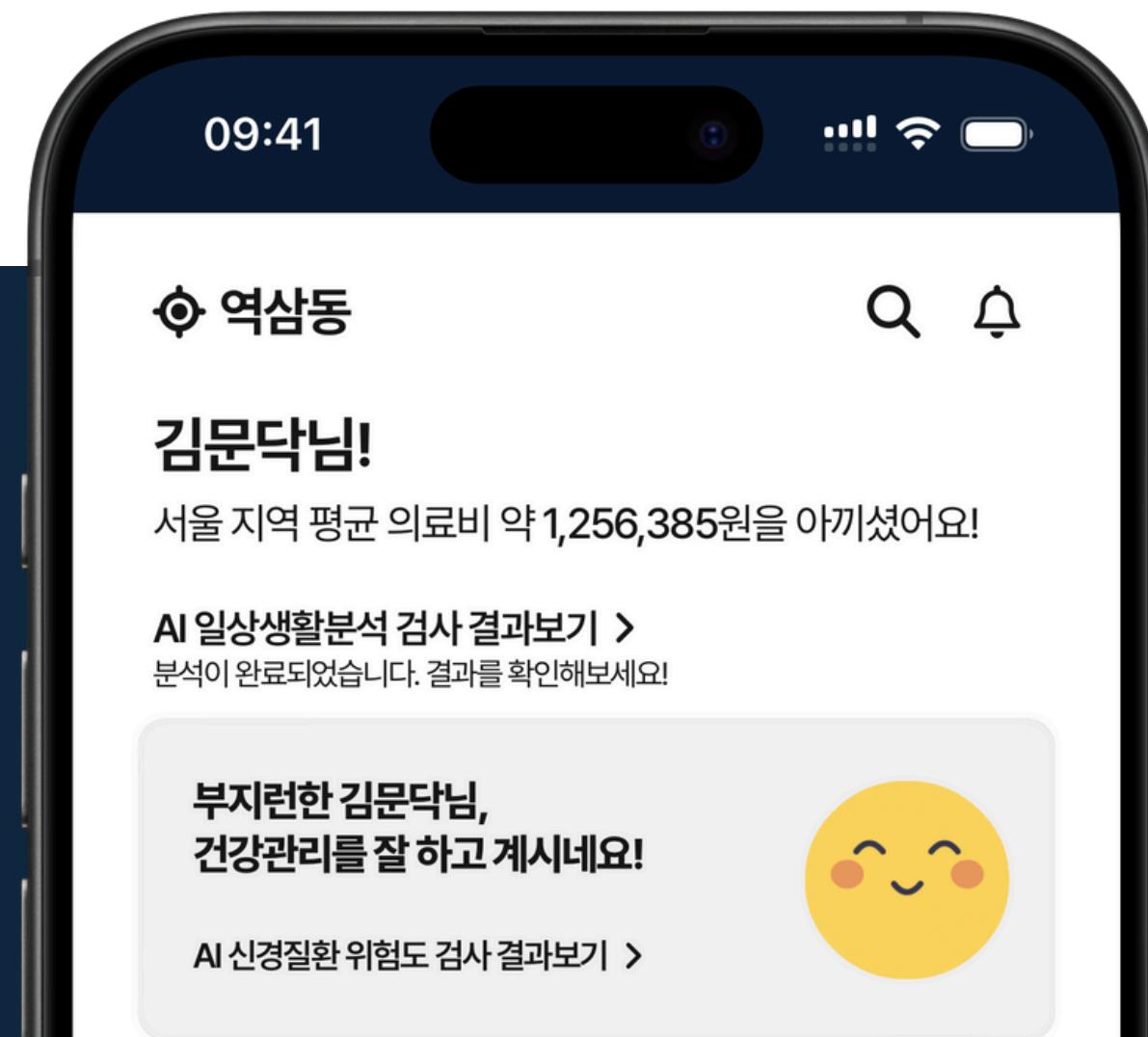
e-index. https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=2766

Hospital newspaper. <https://www.khanews.com/news/articleView.html?idxno=221749>

HIRA. <https://www.hira.or.kr/bbsDummy.do?pgmid=HIRAA020045010000&brdScnBltNo=4&brdBltNo=2399&pageIndex=1&pageIndex2=1>



QnA



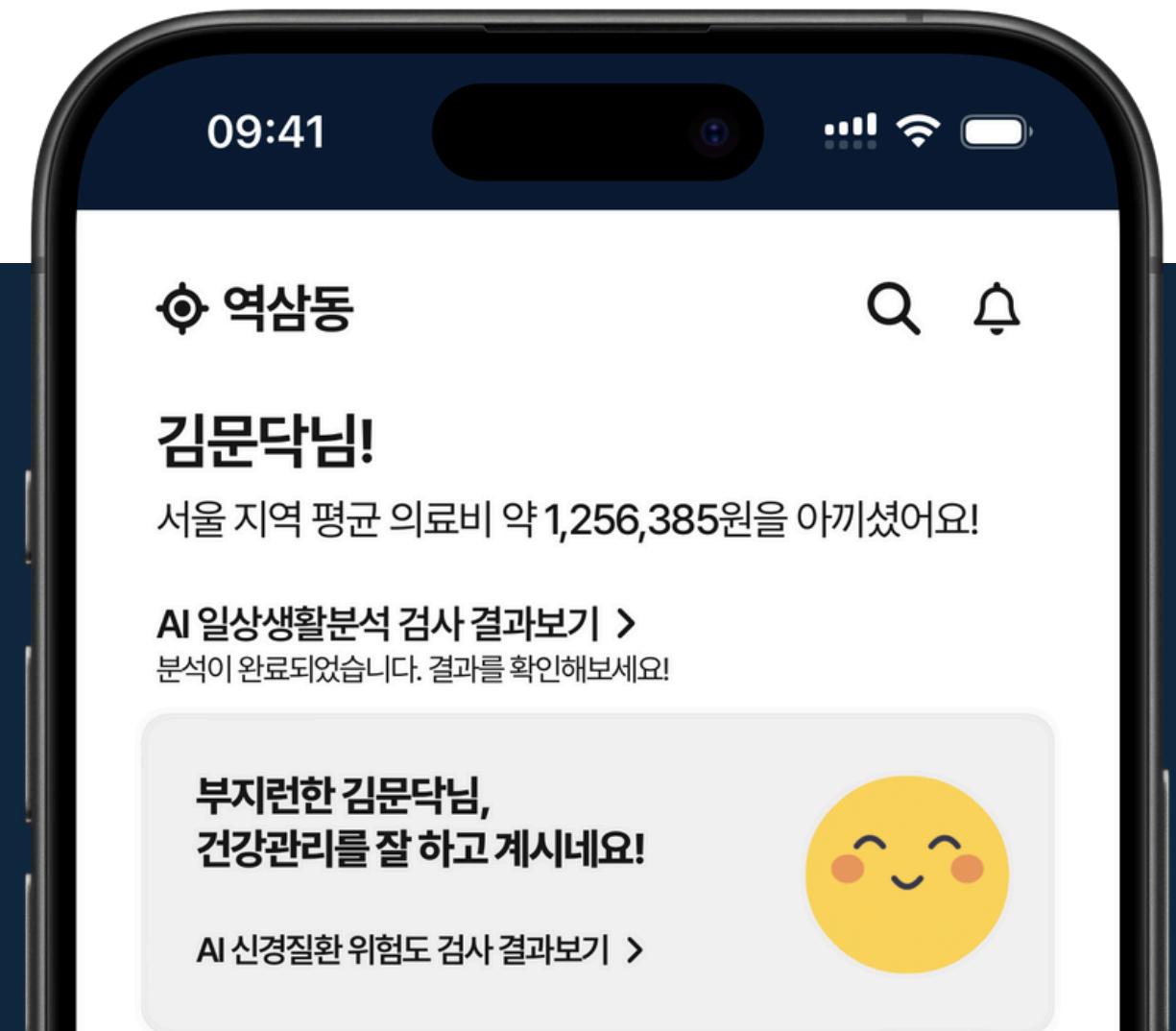
Doctors out the Door

정혜원 박지건 윤지훈 이상훈 임예원 황유진

11-06-2024



Thank you for listening!



Doctors out the Door

정혜원 박지건 윤지훈 이상훈 임예원 황유진

11-06-2024