

메디컬 트윈 기반의 post-TAVR 합병증 예측 모델 연구

발 표 자 : 현 세 민

지도교수 : 이 강 윤



목 차

I. 서론

- 1.1 연구의 배경
- 1.2 연구의 필요성

II. 관련 연구

- 2.1 TAVR
- 2.2 메디컬 트윈
- 2.3 머신러닝 알고리즘

III. 연구 설계

- 3.1 메디컬 트윈 프레임워크 설계
- 3.2 합병증 예측 모델 설계

IV. 실험 및 결과

V. 결론

1. 서론

1.1 연구의 배경

- 인공지능(AI)은 모든 분야에 적용되고 있지만, 특히 헬스케어 분야에서 빠르게 발전하고 있음.
- AI가 헬스케어 분야에 미치는 영향
 - 진료 정확도 향상
 - 치료 계획 최적화
 - 의료 비용 절감
 - 의료 접근성 향상
- AI를 활용한 헬스케어 분야의 주요 응용
 - 임상 의사 결정 지원
 - 약물 개발
 - 의료 영상 분석
 - 개인 맞춤형 의료
 - 합병증 예측을 통한 예방 의료

1.2 연구의 필요성

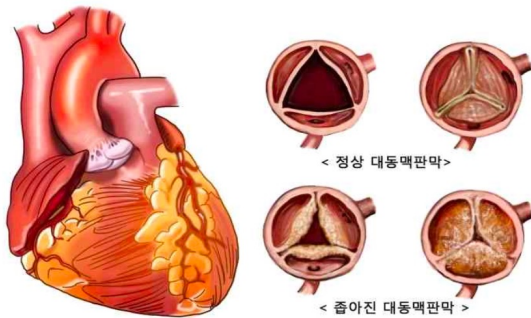
- 합병증이 발생하면 환자의 치료 기간이 길어지고, 치료 비용 증가, 최대 사망까지 이를 수 있음.
- 따라서 합병증을 조기에 예측하고 예방, 완화할 수 있는 방법을 찾는 것이 중요함.
- AI를 활용한 합병증 사례
 - 심장 질환
 - 미국 캘리포니아대학교 버클리 캠퍼스 연구팀 : 심부전 환자의 사망 위험 예측하는 모델 개발
 - 암
 - 미국 스탠퍼드대학교 연구팀 : 암 환자의 재발 위험을 예측하는 모델 개발
 - 수술
 - 미국 MIT 연구팀 : 수술 환자의 감염 위험을 예측하는 모델 개발

2. 관련 연구

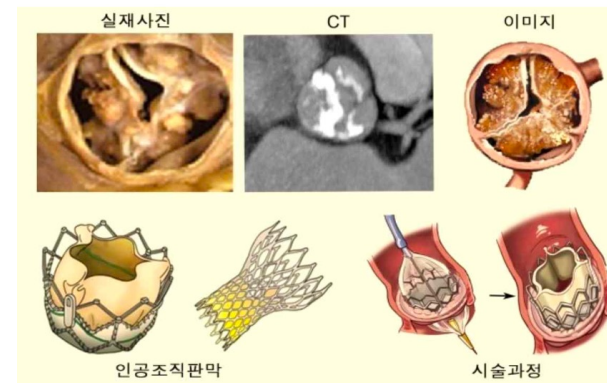
2.1 TAVR

■ 개요

- 대동맥판막은 좌심실과 대동맥 사이에 위치하여 대동맥의 혈액이 좌심실로 역류하는 것을 막아주는 역할을 함.
- 정상적인 대동맥판막은 3 개의 얇은 소엽이 '^'자 모양을 하고 있지만, 선천적 이상 또는 심장 판막에 칼슘이 쌓여 석회화가 진행하면서 발생하는 퇴행성 대 동맥판막 협착증이 발생함.
- 점차 심장 근육 의 비후로 인해 심장 기능 이상으로 호흡곤란, 흉통 및 실신 등의 증상이 나타나고 높은 사망률을 보임.
- 경도관 대동맥판 치환술(TAVR, Transcatheter Aortic Valve Replacement)은 2002년 프랑스 의 Alain Cribier 박사가 중증 대동맥판막협착증 환자에게 세계 최초로 시도해 성공한 이후, 미국과 유럽에서 대규모 임상 연구에서 외과적 판막치환수술(SAVR, Surgical Aortic Valve Replacement)과 비교해서 동등한 성적 을 보여주면서 보편화되고 있음



[그림 1] 대동맥판막 협착증



[그림 2] 경도관 대동맥판 치환술(TVAR)

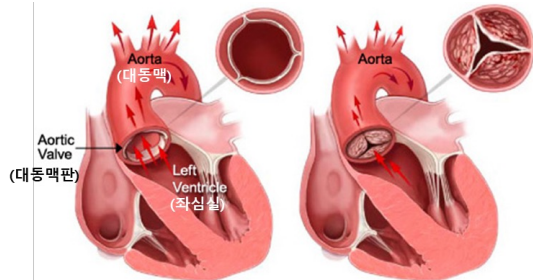
2. 관련 연구

2.1 TAVR

■ 연구 배경

■ 대동맥 협착증 환자의 고령 인구 비율

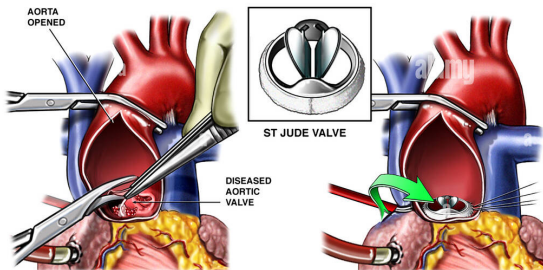
- 70세 이상 노인의 퇴행성 대동맥 협착은 전체 환자의 81.9%를 차지함.



[그림 3] 대동맥 협착증(왼쪽), 정상 대동맥 판막(오른쪽)

■ 수술적 대동맥판 치환술(SAVR)

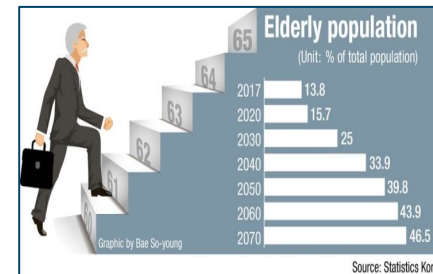
- 중증 대동맥판 환자의 표준 치료는 수술적 대동맥판 치환술이지만 환자의 30-40%는 수술적 치료를 받지 못함



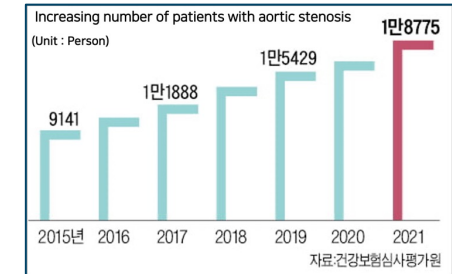
[그림 6] 수술적 대동맥판 치환술(SVAR)

■ 노인인구 증가에 따른 대동맥판 협착증 환자 증가

- 65세 이상 노인인구가 2021년을 기준으로 급증하여 2030년 25% 차지.
- 이에 따라 대동맥판 협착증 환자도 급증할 것으로 예상



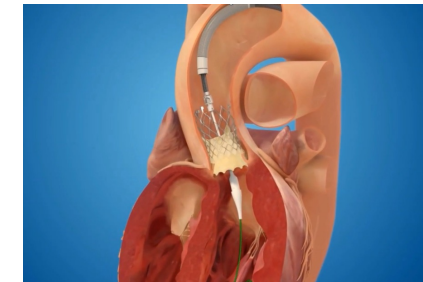
[그림 4] 세계 고령화 인구 변화 그래프



[그림 5] 국내 고령화 인구 변화 그래프

■ 경도관대동맥판치환술(TAVR)의 급속한 증가

- 국내는 연평균 29.4% 증가하여 2021년 3,931건, 2030년 39,986건 시술 예측.
- 해외는 연평균 25% 증가하여 2021년 100,000건, 2030년 745,058건 시술 예측.



[그림 6] 경도관 대동맥판 치환술(TAVR) 과정

2. 관련 연구

2.1 TAVR

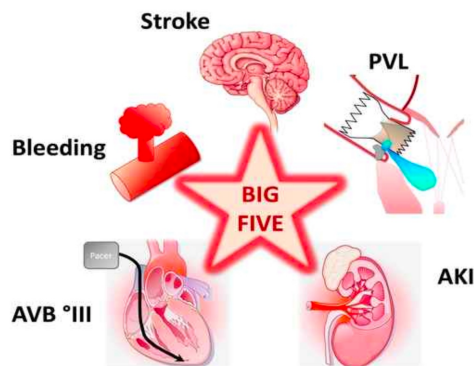
■ TAVR 시술의 문제점

• TAVR 관련 합병증 발생

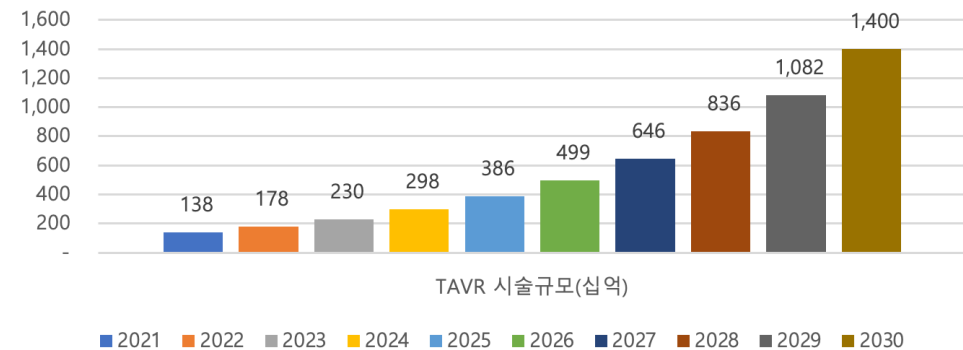
- 시술 후 회복이 빨라서 보편적 치료로 확대 중이지만, 시술이 복잡하고 20~30% 합병증 발생
- 발병 가능한 합병증은 관상동맥 폐쇄, 방실전도 차단, 판막주위 누출, 판막 영구손상 등이 있음.

• 고령화로 인한 TAVR 시술의 급격한 증가로 인한 의료 재정 고갈

- 급속한 노령화로 시술규모가 올해 1780억원에서 2030년 1조4천억원으로 급격히 증가
- 2022년 5월, 수술 불가능군과 고위험군, 80세 이상 중증 대동맥판 협착증 환자에 대한 완전급여화로 의료재정 고갈 심화 우려



[그림 7] TAVR 시술 후 발생 가능한 합병증 BIG FIVE



[그림 8] 국내 TAVR 시술 규모

2. 관련 연구

2.1 TAVR

- 전도 장애(CA; Conduction Abnormalities)
 - 심장의 전기적 신호가 제대로 전달되지 않을 때 발생하며, 불규칙한 심장박동(심부전)이나 심장마비를 일으킬 수 있음.
 - TAVR 후 전도장애가 발생하는 이유는 명확하지 않으나 시술 시 심장의 전기적 신호를 전달하는 신경이 손상되어 발생할 수 있음.
 - 전도장애 발생 시 심장의 전기적 신호를 조절하고 심장의 리듬을 정상화 시킬 수 있는 장치인 심박조율기(pacemaker)를 이식해야 할 수도 있음.
 - 전도 장애에 영향을 미치는 요인은 밸브 크기와 유형, 판막 위치, 기저 질환 등 다양한 요인이 있음.
 - 환자마다 개개인의 특성에 따라 합병증 발병 가능성이 다르므로 개인화된 예측 모델의 필요성 증대.
 - 본 논문에서는 TAVR 후 전도장애를 예측하기 위해 머신러닝 알고리즘을 사용하여 전도장애 예측모델 개발.

2. 관련 연구

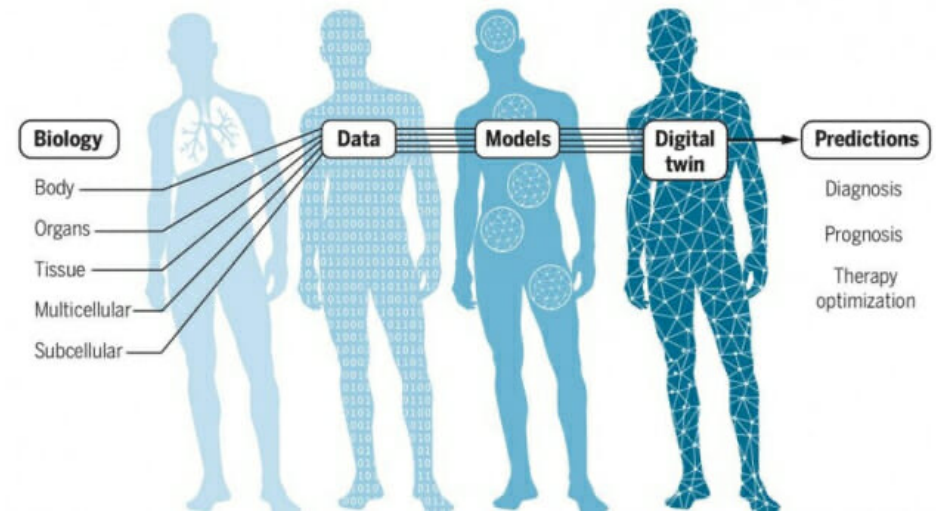
2.2 메디컬트윈

■ 정의

- 메디컬트윈은 실제 환자의 데이터를 기반으로 만들어진 가상의 환자 모델을 의미함.
- 환자의 의료 데이터, 유전 정보, 생활습관 등 다양한 정보를 사용해 생성되며 실제 환자의 상태를 시뮬레이션하고, 치료 계획을 수립하고, 치료 결과를 예측하는 데 활용될 수 있음.

■ 메디컬 트윈 활용 방식

- 질병 예측 및 진단
- 치료 시뮬레이션
- 의료 장비 개발
- 인공지능 및 데이터 분석 훈련



[그림 9] 메디컬 트윈(Medical Twin)

2. 관련 연구

2.2 메디컬트윈

- 장점
 - 치료 계획 최적화
 - 치료 결과 개선
 - 의료 비용 절감

- 메디컬 트윈은 환자의 개인 데이터를 기반으로 생성되며, 의료 분야에서 의사 결정을 지원하고 개인 맞춤형 의료 서비스를 제공하는 역할을 할 수 있음.

2. 관련 연구

2.3 머신러닝 알고리즘

- 로지스틱 회귀(Logistic Regression)

- 통계학과 머신러닝에서 사용되는 분류 알고리즘.
- 분류 문제를 해결하는데 사용됨.
- 주어진 입력 변수들과 그에 따른 결과(이진 분류) 사이의 관계를 모델링하는데 사용되며, 종속 변수가 특정 클래스에 속할 확률을 예측하는데 사용됨.
- 예를 들어, 환자가 특정 질병을 가지고 있는지(1) 없는지(0) 예측함.

$$P(Y = 1 | X) = \frac{1}{1 + e^{-(b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n)}}$$

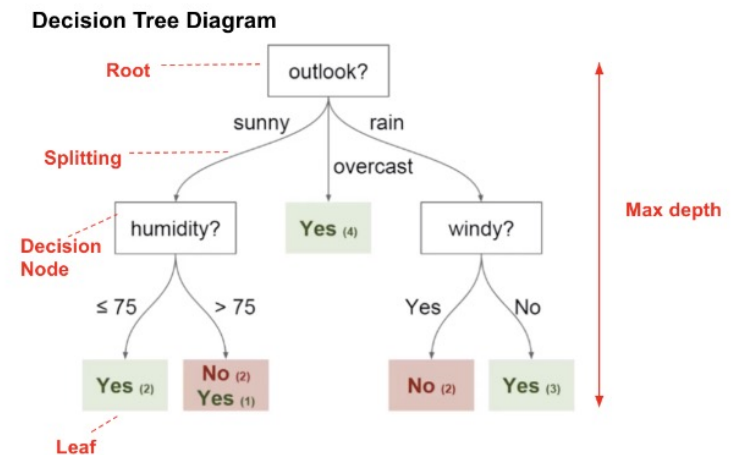
- $P(Y = 1 | X)$: 종속변수 Y 가 1(특정 클래스에 속함)일 확률
- X : 독립변수
- $b_0, b_1, b_2, \dots, b_n$: 회귀 계수

2. 관련 연구

2.3 머신러닝 알고리즘

▪ 의사결정 나무(Decision Tree)

- 학습을 통해 데이터에서 찾은 규칙을 통해 트리 기반 분류 또는 예측을 수행하는 분석 방법
- 즉, 목표변수 범주를 가장 잘 분류하는 예측 변수들로 분류하는 방법
- 루트 노드(Root Node), 의사 결정 노드(Decision Node), 리프 노드(Leaf Node)로 구성
- 규칙 노드는 규칙 조건이 되고, 리프 노드는 결정된 클래스의 값을 나타내고 각각의 새로운 규칙 조건에 대해 서브 트리 생성.
- 데이터 세트의 기능이 규칙 조건을 만들기 위해 서로 결합 할 때마다 규칙 노드가 생성되지만, 너무 많은 규칙은 과적합이 발생할 수 있음. 즉, 트리가 깊어질수록 Decision Tree의 예측 성능이 낮아짐.



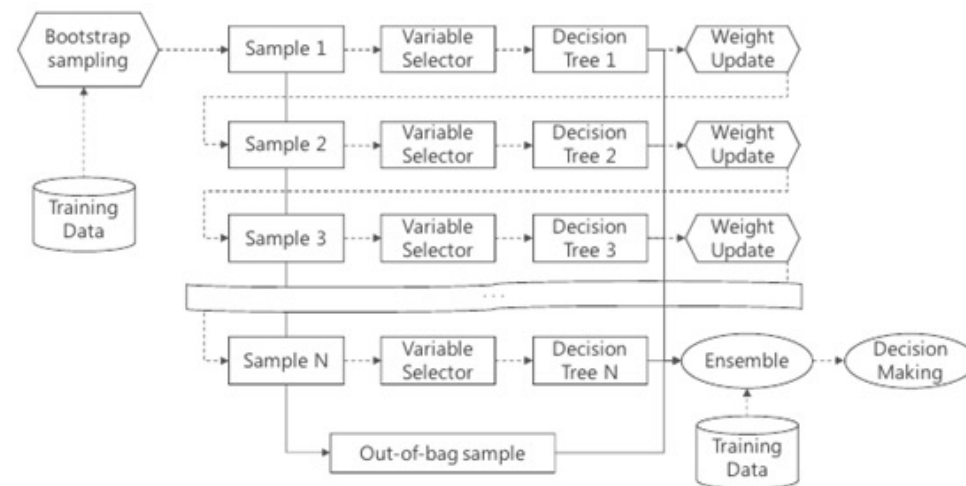
[그림 10] Decision Tree 다이어그램

2. 관련 연구

2.3 머신러닝 알고리즘

■ 랜덤 포레스트(Random Forest)

- 분류와 회귀에 사용되는 지도학습 알고리즘으로, Decision Tree로부터 분류 혹은 회귀 등에서 사용되는 앙상블 학습 기법
- 데이터로부터 부트스트랩 샘플을 추출하여 사용하고, 변수에 랜덤성을 추가하는 방법.
- 앙상블 기법의 장점을 극대화해 기존 방식 대비 분류와 예측의 정확도를 높이는 장점이 있음.
- Out-Of-Bag(OOB) : 부트스트랩 샘플링 과정에서 추출되지 않은 관측치, 주로 테스트 데이터에서의 오분류 비율을 예측하거나 변수 중요도를 추출하는 용도로 사용함.

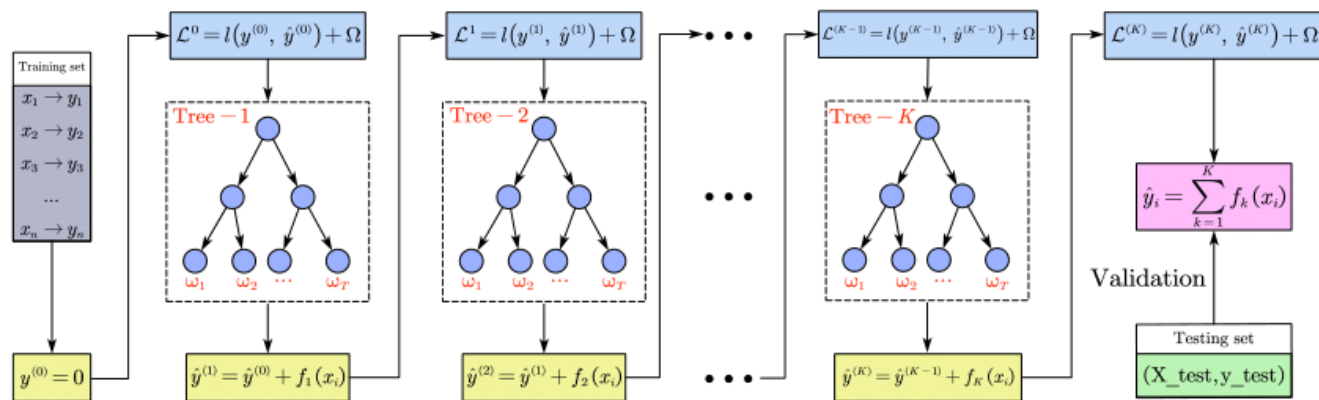


[그림 11] Random Forest 구조도

2. 관련 연구

2.3 머신러닝 알고리즘

- XGBoost (eXtreme Gradient Boosting)
 - 부스팅 기법을 사용해 구현한 Gradient Boost 알고리즘을 병렬학습이 지원되도록 구현한 라이브러리.
 - 회귀와 분류 모두 가능하며, 성능과 자원 효율이 좋음.
 - Gradient Boosting Model 대비 수행시간이 빠르고, XGBoost 자체에 과적합 규제 기능이 있음.
 - 분류와 회귀 영역에서 뛰어난 예측 성능을 발휘하며 조기 종료 기능을 지원함.
 - XGBoost는 손실함수 뿐만 아니라 모형 복잡도까지 고려함.



[그림 12] XGBoost 알고리즘

3. TAVR 메디컬 트윈 프레임워크

3.1 메디컬트윈 프레임워크 설계

■ TAVR 메디컬 트윈 프레임워크

- 현실 세계의 시스템을 수학적 모델로 표현해 특정 조건에서의 동작을 예측하는 환자의 의료 데이터를 기반으로 환자 상태 모니터링, 치료계획을 수립하는 메디컬 트윈의 개념을 활용함.
- 시술 전 · 후의 실세계 데이터(RWD; Real World Data)와 시뮬레이션 데이터가 피드백 시스템을 통해 동기화하여 발전하는 TAVR 메디컬 트윈 프레임워크 연구.



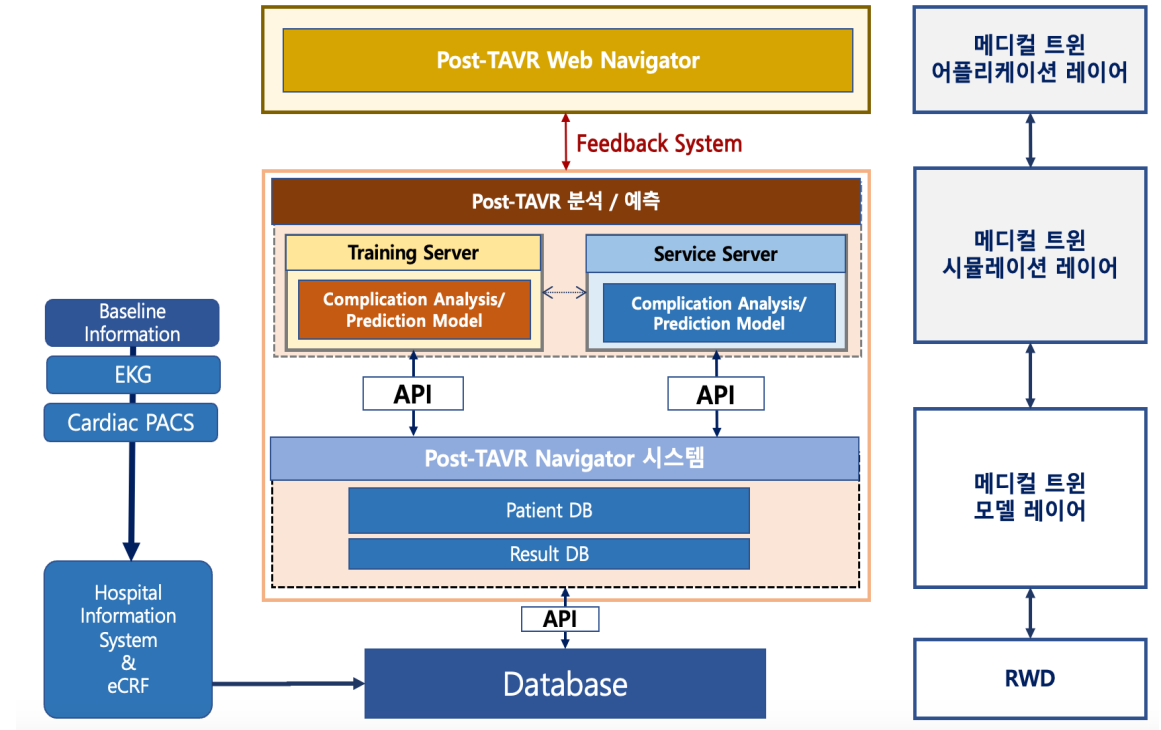
[그림 13] TAVR 메디컬 트윈 프레임워크 레이아웃

3. 연구 설계

3.1 메디컬트윈 프레임워크 설계

■ TAVR 메디컬 트윈 구성도

- RWD
 - 병원에서 수집된 환자 데이터(baseline, EKG, cardiac)는 병원정보시스템(HIS)의 eCRF에 입력되어 데이터베이스에 저장되어 활용됨.
- 메디컬 트윈 모델 레이어
 - RWD는 API를 통해 Post-TAVR Navigator 시스템의 DB에 저장 및 관리.
- 메디컬 트윈 시뮬레이션 레이어
 - Training server와 Service Server에서 데이터 분석 및 합병증 예측모델 개발하고, 결과 내용은 메디컬 트윈 모델 레이어의 ResultDB에 저장함.
- 메디컬 트윈 어플리케이션 레이어
 - Post-TAVR 예측 모델의 결과가 API를 통해 Web Navigator로 전달되고, 웹 대시보드를 통해 분석 내용을 시각화, 예측 모델의 결과에 대한 의료진의 피드백을 진행함.



[그림 14] TAVR 메디컬 트윈 구성도

4. 실험 및 결과

4.1 연구 데이터

- 연구에 사용된 데이터는 인천 길병원에서 연구용으로 제공받은 데이터 활용.
- 변수는 8개의 항목 44개의 변수를 사용함[표1].

Clinic Data	Lab Data	Vital	EKG	Cardiac CT	Cardiac Echo	Drug	Valve
Age	Hemoglobin	sBP	RBBB (Y/N)	LVOT perimeter	AV max	ASA	Valve Code
Gender	HbA1c	dBp	PR interval	LVOT Diameter(mean)	AV mgPG	Antiplatelet other than ASA	Valve Size
BMI	AST	Pulse Rate	First degree AV block		LVEDD	Anticoagulants	
Smoking Status	ALT	Respiratory Rate	QRS interval		LVEF	Beta Blocker	
HTN	Creatinine		Baseline QRS interval > 120ms (Y/N)		Peak AV Vmax	CCB	
DM	Nt-proBNP		Delta PR interval		AV mean pressure gradient	ACEi ARB	
CKD			Delta QRS interval			Diuretics	
			New-onset LBBB (Y/N)			Aldosterone Antagonist	
			Pacemaker Implantation (Y/N)				

[표 1] TAVR 데이터 변수

4. 실험 및 결과

4.2 데이터 탐색

- 범주형 데이터 그룹 간 비교
 - 카이제곱 검정 테스트
- 연속형 변수의 정규 분포
 - 평균 \pm 표준편차(SD)로 표시
 - Student's t테스트를 사용하여 비교
- 연속형 변수의 비정규 분포
 - 중앙값(사분위수 범위)으로 표시
 - Mann-Whitney U 테스트를 사용하여 비교

4.3 실험 설계

- TAVR 후 전도장애 예측 모델 알고리즘
 - Logistic Regression
 - Decision Tree
 - Random Forest
 - XGBoost
- 학습데이터와 테스트데이터는 70:30으로 분할함.
- 결측 데이터 처리는 연속형 변수는 평균값, 이산형 변수는 최빈값으로 대체함.
- 데이터 불균형 문제 해결을 위해 학습 데이터는 SMOTE 기법을 적용함.

4. 실험 및 결과

4.4 결과

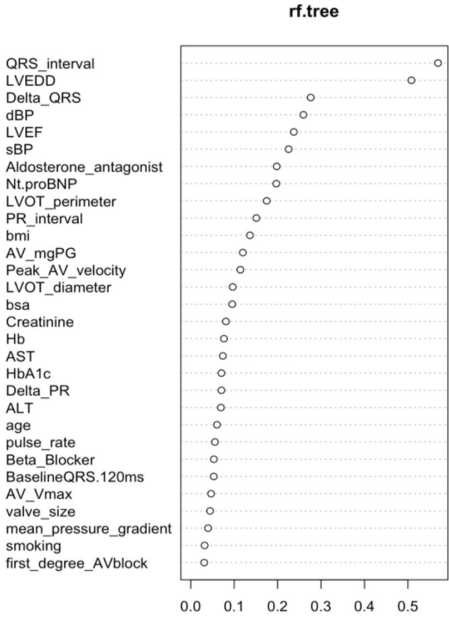
variable	N	Total
age	80	81.0 ± 6.1
Female	43	43 (53.8%)
sBP	80	133.2 ± 15.5
dBP	80	72.6 ± 10.2
RBBB	80	11 (13.8%)
LBBB	80	2 (2.5%)
PR_interval	70	180.7 ± 43.4
first_degree_AVblock	80	11 (13.8%)
QRS_interval	80	99.85 ± 23.2
Delta_PR	65	3.8 ± 40.6
Delta_QRS	80	16.6 ± 26.2
BaselineQRS>120ms	80	14 (17.5%)
New_onset_RBBB	80	7 (8.8%)
New_onset_LBBB	80	11 (13.8%)
pacemaker	80	5 (6.2%)

[표 2] 데이터 기본 특성

	Logistic Regression	Decision Tree	Random Forest	XG Boost
Accuracy	0.83	0.87	0.90	0.83
F1-score	0.33	0.57	0.56	0.33
AUC	0.62	0.70	0.50	0.60

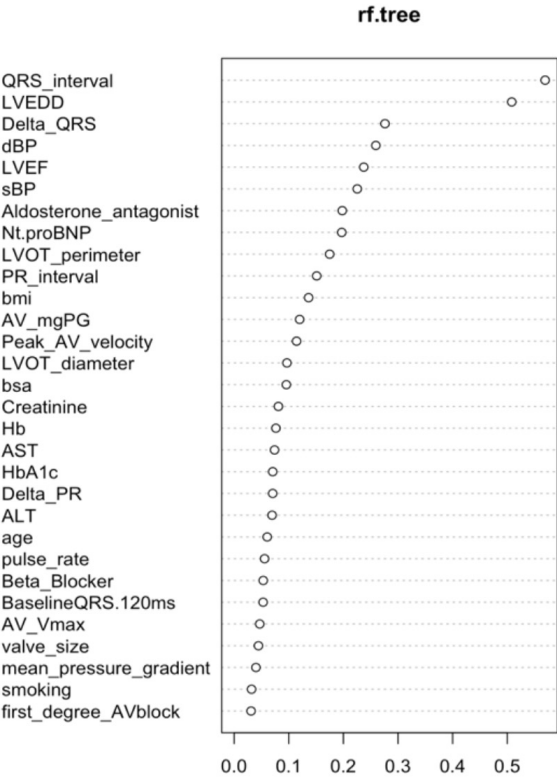
[표 3] 각 알고리즘 정확도 비교

[그림 15] ROC Curve

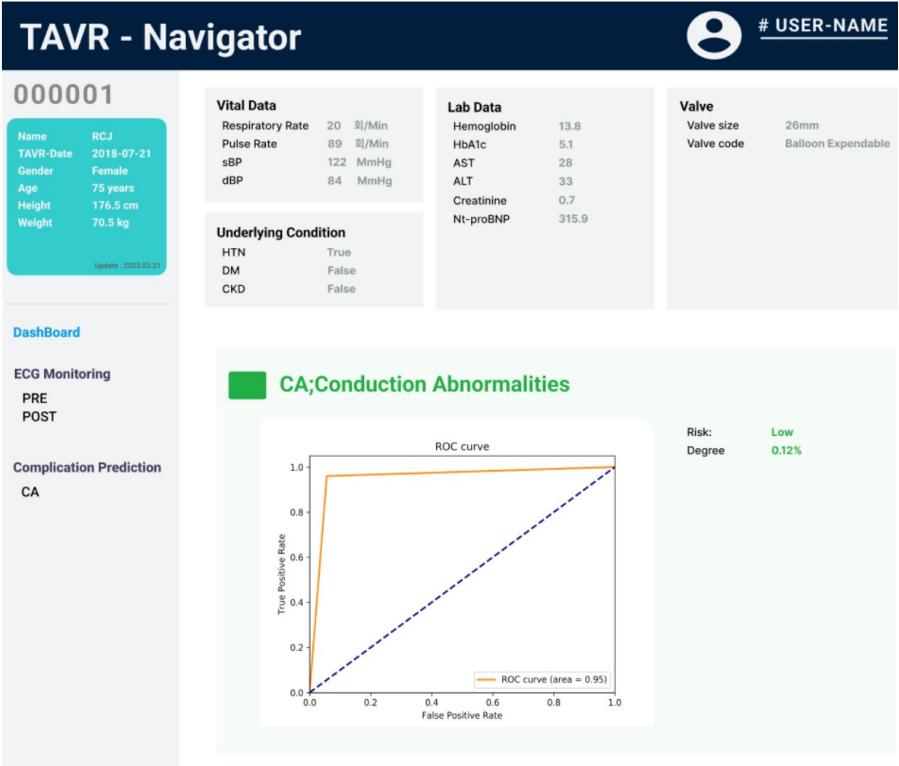


4. 실험 및 결과

4.4 결과



[그림 16] ROC Curve



[그림 17] TAVR Web Navigator(TWN)

5. 결론

- 본 논문에서는 다양한 의료 시술 중 경도관 대동맥판 치환술(TAVR)를 선정하여 메디컬 트윈 프레임워크를 구성하고 TAVR 후 발생 가능한 합병증 중 하나인 전도장애를 머신러닝 알고리즘을 통해 예측모델을 개발함.
- TAVR 후 발생 가능한 전도장애의 경우, ECG(심전도)가 다른 항목들에 비해 영향을 미치는 것을 파악함.
- 즉, 메디컬 트윈 프레임워크를 기반으로 RWD 데이터를 활용해 시술 후 합병증 예측을 진행하고, 해당 결과를 웹 대시보드를 통해 분석보고서를 제공하고, 의료진의 피드백을 통해 정확도 개선함으로써 보다 나은 서비스를 제공할 수 있을 것으로 기대됨.
- 현실세계에서 분석 및 예측 시뮬레이션을 진행했던 기존 연구와 비교하여 Medical Twin은 데이터 수집에서 피드백에 이르기까지 TAVR의 전 과정이 가상 환경에서 이루어질 수 있도록 하여 시술 전 진단 및 평가 과정에서 의사가 객관적이고 구체적인 데이터를 기반으로 수술 결과를 예측하고 임상 결과를 개선할 수 있도록 하며, 편의성 및 개선 효과가 기대됨.
- 장기적으로 의료 빅데이터 구축을 통해 국가의료정보화에 기여하고, 재시술이나 합병증 감소로 인한 보건 의료비용 감소, 환자의 경우 삶의 질 향상과 노인 환자 의료비 부담 감소, 건강한 기대수명 연장으로 환자 삶의 만족도 향상에 기여할 것으로 기대됨.

Gachon
Cognitive Computing Lab

