

2021 데이터 청년 캠퍼스 경기대학교

C/O 리스크 최소화를 위한 심각도 예측 모델

김덕일, 박하림, 신승균, 최운서, 황시연

목차

①

프로젝트 배경

4p

②

데이터 설명

5p

③

데이터 전처리

9p

④

모델링

15p

⑤

R shiny 구현

22p

⑥

기대효과

23p

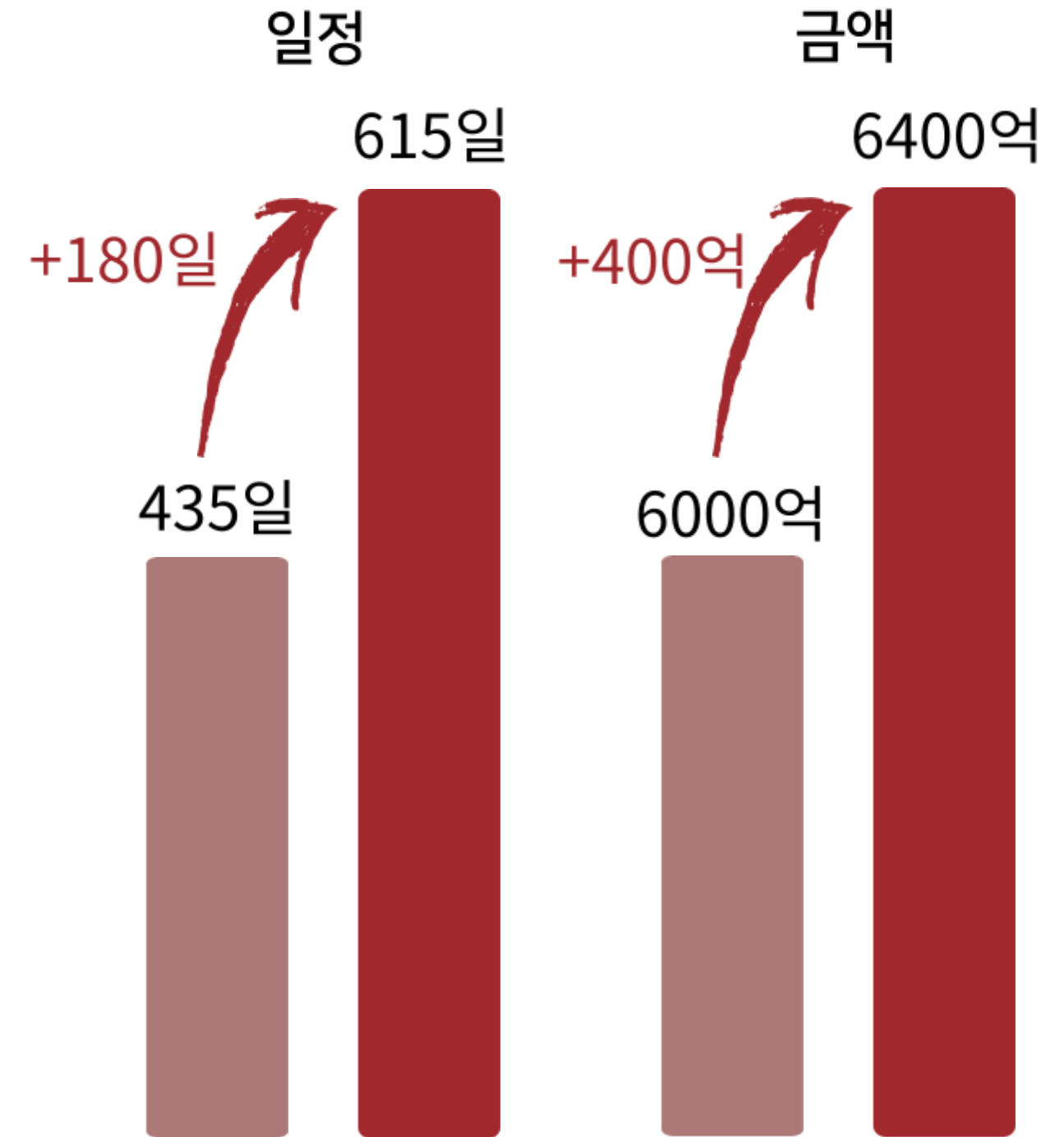
프로젝트 배경



<플랜트 공정 과정과 c/o발생 구역>

C/O (Change Order)

- 플랜트공정 중 설계변경 요구
- 플랜트 project 당 2~3 만건 발생
- 시간, 비용적인 손해 발생시 시공사측의 책임



<한 플랜트 공정 당 평균 일정, 금액 증가 수치>

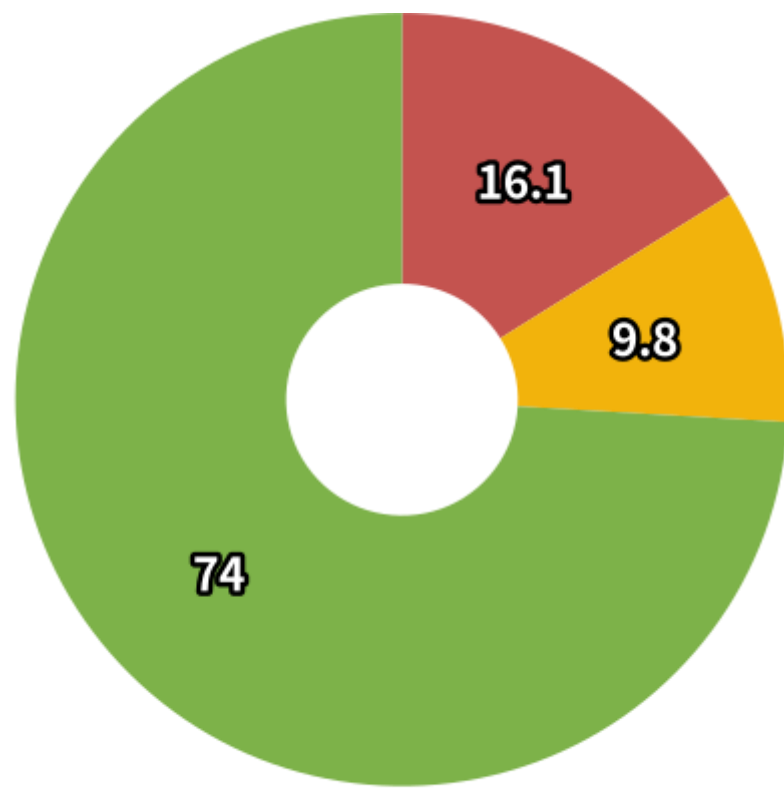
데이터설명

1029개 행 X 34개 열

프로젝트 SEQ	프로젝트 명칭	프로젝트 코드	프로젝트 분야	플랜트 종류	유형	규모/용량	Location	총 공사비	공사기간	발주처	착수일	설계비용	설계기간
1	KGHT	16010200001	화공	육상플랜트	Refinery	염화에틸렌 30만톤을 생산	DUBAI	95,900,000,000	38	SABIC	2016-01-02	6,138,000,000	480
2	YAC	13080100002	화공	육상플랜트	Refinery	30만톤 규모의 저밀도 폴리에틸렌 생산	SAUDI	9,850,000,000	36	ARAMCO	2013-08-01	608,000,000	480
15	JSC	11080100015	화공	육상플랜트	Refinery	59,000 ton	SAUDI	22,300,000,000	30	Saudi Aramco	2011-08-01	1,402,500,000	420
27	KDK	10040100027	부유식	해상플랜트	부유식	Drilling Depth: 20, 100ft	카자흐스탄	19,000,000,000	27	국영석유회사	2010-04-01	1,207,375,000	420

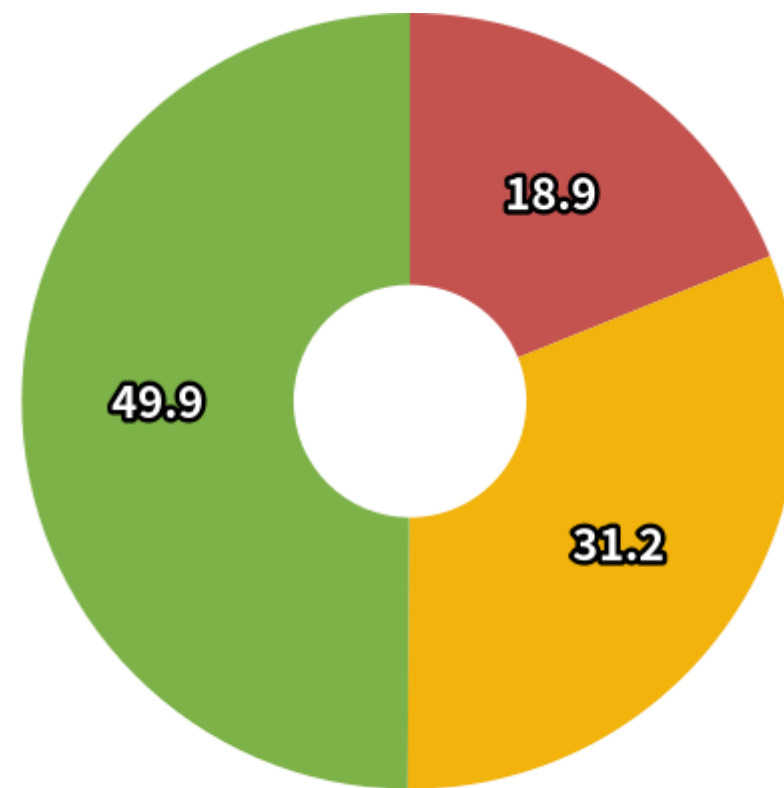
설계변경 공종	설계변경 유형	설계변경 사유	Code	M/H	투입인원	변경착수일	일정심각도	설계변경 기간	일정(%)	금액심각도	투입금액	금액(%)
배관	DWG 추가	Total ISO. DWG : 12030 Sheet의 10% (Sheets)	KGHT_0001	1860	14	2016-03-06	심각	17	3.46%	심각	124,406,100	2.03%
기계	VALVE 추가	*3"-LS-UY10018-NA1S-PPI - VALVE 추가	YAC_0080	3	1	2014-08-04	안전	0	0.08%	안전	245,700	0.04%
건축	사업주 요청사항	Hoist Beam 추가 (15 EA)	JSC_0015	458	5	2013-07-09	심각	11	2.73%	심각	30,633,330	2.18%
전기	Location 변경	VSAT Antenna는 공간 확보가 용이한 공기압축기 Module 상부로 배치하는 방안 검토	KDK_0002	270	5	2010-09-27	경계	7	1.61%	경계	22,113,000	1.83%

일정심각도



■ 심각 ■ 경계 ■ 안전

금액심각도



■ 심각 ■ 경계 ■ 안전

<일정심각도(좌), 금액심각도(우) 데이터 내 분포>

심각도 (종속변수)

플랜트공정설계 중 설계변경에 있어서
기간이나 비용에 **얼마나 타격을 주는지**에 대한 지표

- 일정심각도(%) = 설계변경기간(일) / 설계기간(일)
- 금액심각도(%) = 설계변경금액(원) / 설계금액(원)

* 범주형 변수로 변환시 기준

- 1% 미만 : 안전
- 1~2 % : 경계
- 2% 이상 : 심각

데이터설명

1029개 행 X 34개 열

프로젝트 SEQ	프로젝트 명칭	프로젝트 코드	프로젝트 분야	플랜트 종류	유형	규모/용량	Location	총 공사비	공사기간	발주처	착수일	설계비용	설계기간
1	KGHT	16010200001	화공	육상플랜트	Refinery	염화에틸렌 30만톤을 생산	DUBAI	95,900,000,000	38	SABIC	2016-01-02	6,138,000,000	480
2	YAC	13080100002	화공	육상플랜트	Refinery	30만톤 규모의 저밀도 폴리에틸렌 생산	SAUDI	9,850,000,000	36	ARAMCO	2013-08-01	608,000,000	480
15	JSC	11080100015	화공	육상플랜트	Refinery	59,000 ton	SAUDI	22,300,000,000	30	Saudi Aramco	2011-08-01	1,402,500,000	420
27	KDK	10040100027	부유식	해상플랜트	부유식	Drilling Depth: 20, 100ft	카자흐스 탄	19,000,000,000	27	국영석유 회사	2010-04-01	1,207,375,000	420

설계변경 공종	설계변경 유형	설계변경 사유	Code	M/H	투입인원	변경착수일	일정심각도	설계변경 기간	일정(%)	금액심각도	투입금액	금액(%)
배관	DWG 추가	Total ISO. DWG : 12030 Sheet의 10% (Sheets)	KGHT_0001	1860	14	2016-03-06	심각	17	3.46%	심각	124,406,100	2.03%
기계	VALVE 추가	*3"-LS-UY10018-NA1S-PPI - VALVE 추가	YAC_0080	3	1	2014-08-04	안전	0	0.08%	안전	245,700	0.04%
건축	사업주 요청사항	Hoist Beam 추가 (15 EA)	JSC_0015	458	5	2013-07-09	심각	11	2.73%	심각	30,633,330	2.18%
전기	Location 변경	VSAT Antenna는 공간 확보가 용이한 공기압축기 Module 상부로 배치하는 방안 검토	KDK_0002	270	5	2010-09-27	경계	7	1.61%	경계	22,113,000	1.83%

프로젝트 분야	플랜트 종류	Location	발주처	설계변경 공종
화공	육상플랜트	DUBAI	SABIC	배관
화공	육상플랜트	SAUDI	ARAMCO	기계
화공	육상플랜트	SAUDI	Saudi Aramco	건축
부유식	해상플랜트	카자흐스탄	국영석유 회사	전기
설계변경 사유				
Total ISO. DWG : 12030 Sheet의 10% (Sheets)				
*3"-LS-UY10018-NA1S-PPI - VALVE 추가				
Hoist Beam 추가 (15 EA)				
VSAT Antenna는 공간 확보가 용이한 공기압축기 Module 상부로 배치하는 방안 검토				

독립변수

- 프로젝트 분야 : 플랜트 공정 분야

levels : 화공, 고정식, 부유식, 발전

- 플랜트 종류 : 플랜트 종류

levels : 육상플랜트, 해상플랜트

- Location : 공사 진행 지역

levels : DUBAI, 한국, 카자흐스탄, ...

- 발주처 : 발주 기관

levels : SABIC, 국영석유회사, ...

- 설계변경공종 : 설계 중 변경요청된 분야

levels : 배관, 기계, 전기, ...

- 설계변경사유 : 설계변경이 발생한 사유 (Text Data)

<설계변경사유 텍스트 데이터 전처리 개요>

설계변경 사유
Total ISO. DWG : 12030 Sheets의 10% (Sheets)
P&ID 추가 및 변경에 따른 해석대상 추가
계약시 383 Lines -----> 575 Lines으로 증가.
계약시 300 Item ----->491 Item으로 증가
Reactor structure foundation IFC drawing issue를 위하여 PMT 응력해석 scope인 reaction loop 9 system
Model Comment 중 사업주 Comment에의한 C/O 사항.
Add Work (57 Lines)
Add Work (190Lines)
첨부 설계 변경 요청사항중 추가 M/H 요청 사항(1명).
Desuper Heater (DS1/2-101101 & DS1/2-101102)변경에 따른 설계 수정.
계장 valve에 대한 Final Information에의거 Cata DB 변경에 따라 3D Model Up-Date Control valve 변경 (
적용 대상 equipment 증가.
NS-1 Pipe Rack 동쪽에 설치되는 Pump의
설계 변경및 3d model 수정.
Line size / class변경.
Feed 대비 Utility Station & Safety Shower의
사업주 요청에 의한 변경 작업. · P&ID 변경
Project Schedule 53m -----> 53m.
다수 16명 -----> 98명 / 현역 11명
Project Acceleration을 위한 over time 및 휴일근무에 따른 M/H 소진.16명 x 3시간 / day x 25일 x 3 Mont
Line 증가에 따른 업무량 증가.
Line 증가에 따른 업무량 증가.
Geo.Dwg Total : 5527 Sheets
계약시 940 Sheets -----> 9904 Sheets으로 증가
업무량 증가에 따른 Plan Dwg장수 증가.
Line추가에 따른 ISO. DWG출도 관련 M/H증가
계약 시 12030 Sheets -----> 13433 Sheets
Equip. Modeling 추가 1) P-101132A/B (OILY WATER PIT PUMPS) 2) P-101133A/B (OILY WATER PI
계약시 3800 Lines ----> 4601 Lines
고압가스법에 따른 이격거리 바뀜에 의거 Plot Plan의 변경

<전처리 전 데이터 일부>



	P&ID	dwg	...	변경	추가
1	1	1		0	0
2	0	0		1	0
...					
1029	1	1		0	1

<전처리 후 데이터 형태 - 이진행렬>

<설계변경사유 Data 탐색>

원데이터	포개진 단어	문제점
설계 변경및 3d model 수정	설계, 변경및, 3d, model, 수정	조사 삭제 필요
Line추가에 따른 Plan Dwg장수 증가	Line추가에, 따른, Plan, Dwg장수, 증가	띄어쓰기 필요
VENDOR PRINT 변경으로 LOCATION 변경	VENDOR, PRINT, 변경으로, LOCATION, 변경	도메인 단어 확인 필요
400-31A30-MDS-CA09-0110-NI, 2-BRANCH 추가	400-31A30-MDS-CA09-0110-NI, , 2-BRANCH, 추가	코드 삭제 필요
계약시 210 lines ---> 400 lines으로 증가.	계약시, 210, lines, --->, 400, lines으로, 증가.	단어 정형화 필요
line 증가에 따른 업무양 증가.	line, 증가에, 따른, 업무양, 증가.	
lc : 29 라인 증가	lc, :, 29, 라인, 증가	

데이터 전처리

1. 단어 정형화

1-1. 플랜트 사전 구축

- 도메인 지식을 활용한 단어 묶기 작업 필요
- ex) plot-plan, by-pass, vender-print, change-order

1-2. 소문자화

- 위에서 만든 플랜트 사전(단어 변경 파일)을 적용시키기 위해 소문자화
- ex) VENDER -> vender, LOCATION -> location

1-3. 동의어 통일

- 같은 의미를 갖지만 다른 형태를 한 단어 통일
- ex) line, lines, 라인 -> line
- drawing -> dwg

변경전	변경후	
plot plan	plot-plan	
펌프	pump	
pumps	pump	
ug	u/g	
vp	v/p	
3d model	3d-model	
air compres	air-compressor	
air comp	air-compressor	
analzer po	analzer-point	
by pass	by-pass	
데크	deck	
디자인	design	
desuper h	desuper-heater	
ewatering	dewatering	
탈수	dewatering	

<단어변경.xlsx>

	P&ID	dwg	...	변경	추가
1	1	1		0	0
2	0	0		1	0
...					
1029	1	1		0	1

<전처리 후 데이터 형태 - 이진행렬>

전처리 전/후 오분류표 및 정확도 비교

- 예측변수: 설계변경사유로 만든 이진행렬
- 종속변수: 일정심각도
- 사용 모델: 나이브 베이즈

	안전	경계	심각
안전	150	5	3
경계	10	6	2
심각	29	3	4

VS

	안전	경계	심각
안전	134	9	7
경계	9	8	3
심각	10	6	20

검증데이터 정확도: 0.7547

검증데이터 정확도: 0.7864

<전처리 전>

<전처리 후>

데이터 전처리

범주형 독립변수					논리형 독립변수					종속변수	
프로젝트분야	플랜트종류	Location	발주처	설계변경공종	P&ID	dwg	...	변경	추가	일정심각도	금액심각도
화공	육상 플랜트	DUBAI	SABIC	배관	0	0	0	0	0	경계	안전
고정식	육상 플랜트	DUBAI	SABIC	배관	0	0	0	0	0	안전	안전
화공	육상 플랜트	Kuwait	ARAMCO	배관	0	0	0	0	0	안전	안전
발전	육상 플랜트	Kuwait	ARAMCO	배관	0	0	0	0	0	안전	안전
고정식	해양 플랜트	한국	Duqm	배관	0	0	0	0	0	경계	안전
화공	해양 플랜트	한국	Duqm	배관	0	0	0	0	0	안전	안전
...
부유식	해양 플랜트	카자흐스탄	국영석유회사	배관	0	0	0	0	0	심각	경계
발전	해양 플랜트	카자흐스탄	국영석유회사	배관	0	0	0	0	0	심각	심각
화공	해양 플랜트	카자흐스탄	국영석유회사	배관	0	0	0	0	0	심각	심각
고정식	육상 플랜트	SAUDI	ARAMCO	배관	0	0	0	0	0	심각	경계
...
화공	육상 플랜트	SAUDI	ARAMCO	배관	0	0	0	0	0	심각	경계
고정식	육상 플랜트	태국	Petrofac	배관	0	0	0	0	0	심각	경계

범주형 독립변수

논리형 독립변수

종속변수

훈련 검증 데이터 분리

프로젝트분야	플랜트종류	Location	발주처	설계변경공종	P&ID	dwg	...	변경	추가	일정심각도	금액심각도
화공	육상 플랜트	DUBAI	SABIC	배관	0	0	0	0	0	경계	안전
고정식	육상 플랜트	DUBAI	SABIC	배관	0	0	0	0	0	안전	안전
화공	육상 플랜트	Kuwait	ARAMCO	배관	0	0	0	0	0	안전	안전
발전	육상 플랜트	Kuwait	ARAMCO	배관	0	0	0	0	0	안전	안전
고정식	해양 플랜트	한국	Duqm	배관	0	0	0	0	0	경계	안전
화공	해양 플랜트	한국	Duqm	배관	0	0	0	0	0	안전	안전
...
부유식	해양 플랜트	카자흐스탄	국영석유회사	배관	0	0	0	0	0	심각	경계
발전	해양 플랜트	카자흐스탄	국영석유회사	배관	0	0	0	0	0	심각	심각
화공	해양 플랜트	카자흐스탄	국영석유회사	배관	0	0	0	0	0	심각	심각
고정식	육상 플랜트	SAUDI	ARAMCO	배관	0	0	0	0	0	심각	경계
...
화공	육상 플랜트	SAUDI	ARAMCO	배관	0	0	0	0	0	심각	경계
고정식	육상 플랜트	태국	Petrofac	배관	0	0	0	0	0	심각	경계

훈련데이터
(80%)

검증데이터
(20%)

분석 과정

방법 1) 수치형 종속변수 예측 후 범주형으로 변환

프로젝트분야	플랜트종류	...	P&ID	dwg	...	추가
화공	육상플랜트		0	1		0
고정식	육상플랜트		0	0		0
화공	육상플랜트		1	0		0
발전	육상플랜트		1	1		0
고정식	해양플랜트		0	1		1
...			
화공	해양플랜트		0	1		1
...			
부유식	해양플랜트		0	0		1
발전	해양플랜트		1	0		0
화공	해양플랜트		1	1		0
고정식	육상플랜트		0	0		0

선형회귀, SVM

심각도(수치)
3.73%
0.5%
0.26%
1.78%
1.88%
...
2.56%
...
0.89%
2.56%
1.45%
0.23%

범주형으로 변환

심각도(범주)
심각
안전
안전
경계
경계
...
심각
...
안전
심각
경계
안전

→ 예측 정확도: 50 ~ 60% (X)

→ 범주형 종속변수 예측(방법2,3)

분석 과정

방법 2) 설계변경사유가 **설계변경유형으로 분류될 확률값**을 예측변수로 사용

	설계변경 유형	설계변경 사유
1	DWG 추가	Total ISO. DWG : 12030 Sheets의 10% (Sheets)
2	P&ID 변경	"P&ID 추가 및 변경에 따른 해석대상 추가 계약시 210 Lines -----> 400 Lines으로 증가."
3	Line 추가	계약시 383 Lines -----> 575 Lines으로 증가.
4	Line 추가	계약시 300 Item ----->491 Item으로 증가
...
1027	사업주 요청사항	Valve Access 할 수 있게 공간 확보(3 DWG)
1028	사업주 요청사항	Elevator 고장 대비 Crane으로 Bag 놓을 수 있게 Platf orm 연장(12 DWG)
1029	사업주 요청사항	PPB 접근을 위해 House를 세로 배치 고려하여 도로에 서 접근할 수 있도록 수정((3 DWG)

나이프베이즈

	DWG 추가	P&ID 변경	Line추가	...	Spec추가	...
1	0.9938	0.0002	3.44E-10	...	0.0059	...
2	0.0013	8.22E-05	3.76E-07	...	0.0119	...
3	0.0242	2.29E-05	1.03E-05	...	0.0025	...
4	0.2786	0.0002	1.21E-05	...	0.0167	...
...
1027	8.53E-05	0.7579	7.17E-07	...	0.2406	...
1028	0.0083	0.0037	2.07E-06	...	0.4488	...
1029	0.0015	0.8456	1.57E-06	...	0.1498	...

예측

→ 학습데이터 정확도: **0.9745**, 검증데이터 정확도: **0.8350** → 과적합

*랜덤포레스트 기준

분석 과정

방법 3) 기존 전처리된 데이터 형태 그대로 범주형 종속변수 예측

프로젝트분야	플랜트종류	...	P&ID	dwg	...	추가
화공	육상플랜트		0	1		0
고정식	육상플랜트		0	0		0
화공	육상플랜트		1	0		0
발전	육상플랜트		1	1		0
고정식	해양플랜트		0	1		1
...			
화공	해양플랜트		0	1		1
...			
부유식	해양플랜트		0	0		1
발전	해양플랜트		1	0		0
화공	해양플랜트		1	1		0
고정식	육상플랜트		0	0		0

랜덤포레스트, 나이브베이지, KNN, 부스팅, 다항회귀

심각도(범주)
심각
안전
안전
경계
경계
...
심각
...
안전
심각
경계
안전

➔ 학습데이터 정확도: 0.9356, 검증데이터 정확도: 0.8544 *랜덤포레스트 기준

방법2, 방법3 비교

(방법2) 예측변수 변형 0

	안전	경계	심각
안전	607	0	2
경계	10	69	3
심각	2	4	126

학습데이터 정확도

0.9745

<방법2 일정심각도 오분류표 및 정확도>

	안전	경계	심각
안전	144	3	3
경계	5	10	5
심각	11	7	18

검증데이터 정확도

0.8350

VS

(방법3) 예측변수 변형 X

	안전	경계	심각
안전	608	1	0
경계	26	53	3
심각	20	3	109

학습데이터 정확도

0.9356

<방법3 일정심각도 오분류표 및 정확도>

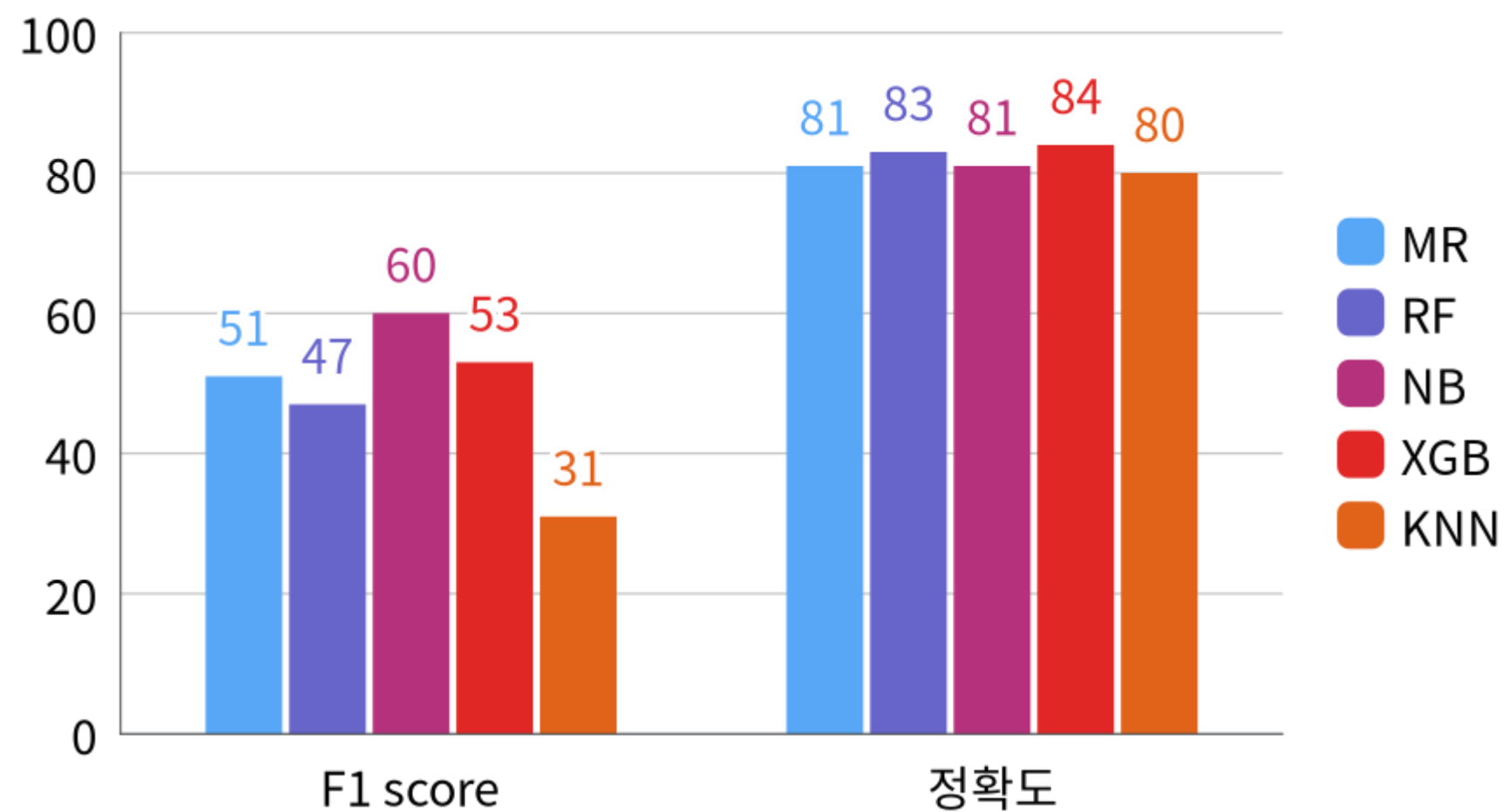
	안전	경계	심각
안전	146	0	4
경계	6	9	5
심각	12	3	21

검증데이터 정확도

0.8544

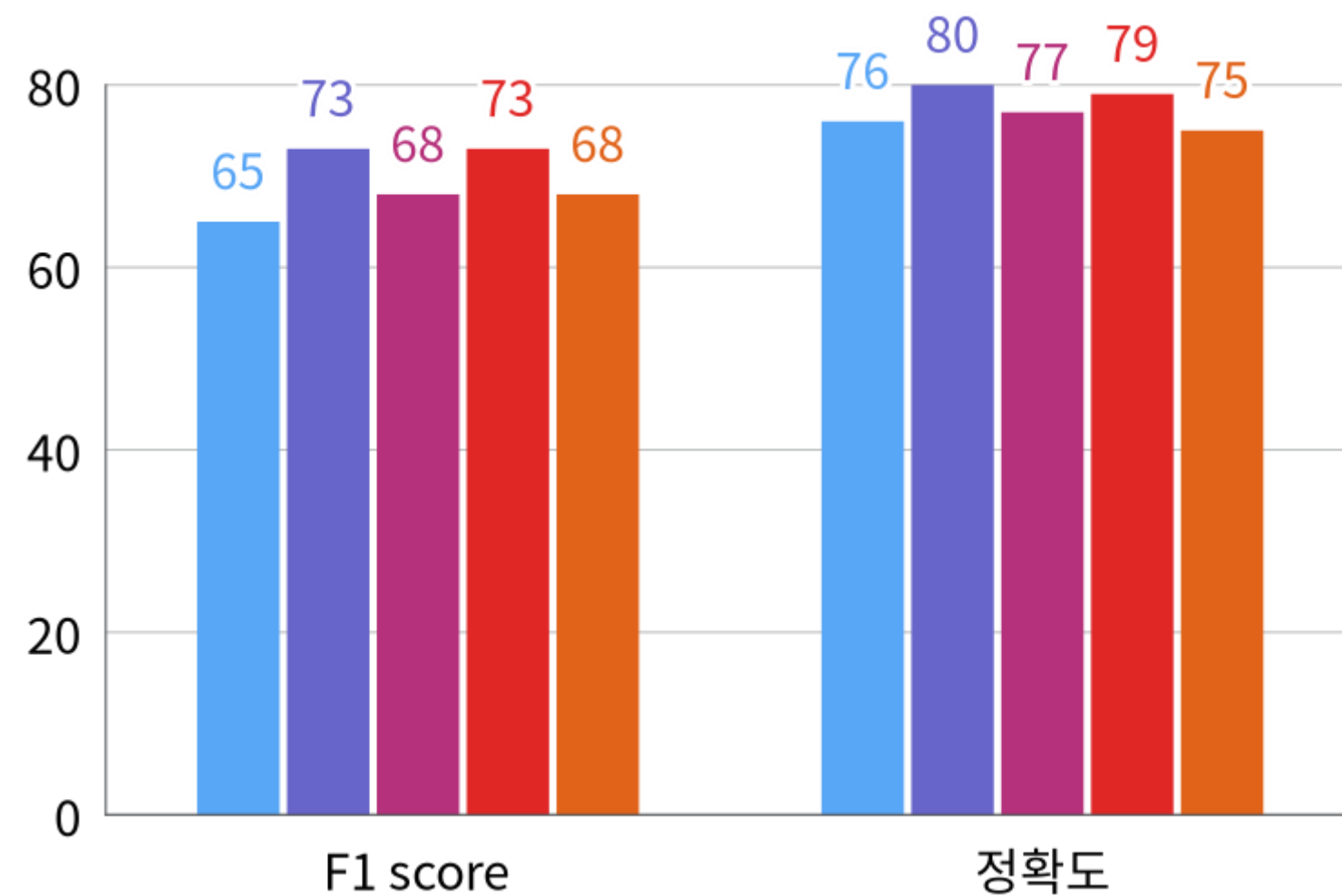
방법3 모델링 결과 비교

<일정 심각도>



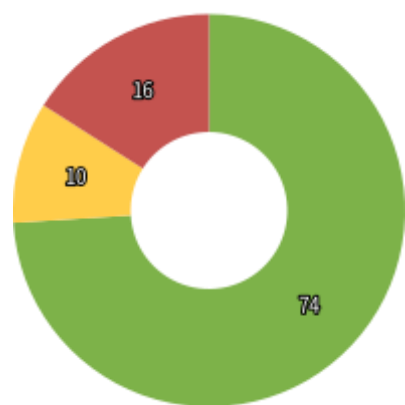
	MR	RF	NB	XGB	KNN
F1 score	51.16	47.05	60.46	53.33	31.25
정확도	81.55	83.49	81.55	84.46	50.08

<금액 심각도>



	MR	RF	NB	XGB	KNN
F1 score	65.64	73.84	68.8	73.5	68.25
정확도	76.69	80.09	77.18	79.12	75.24

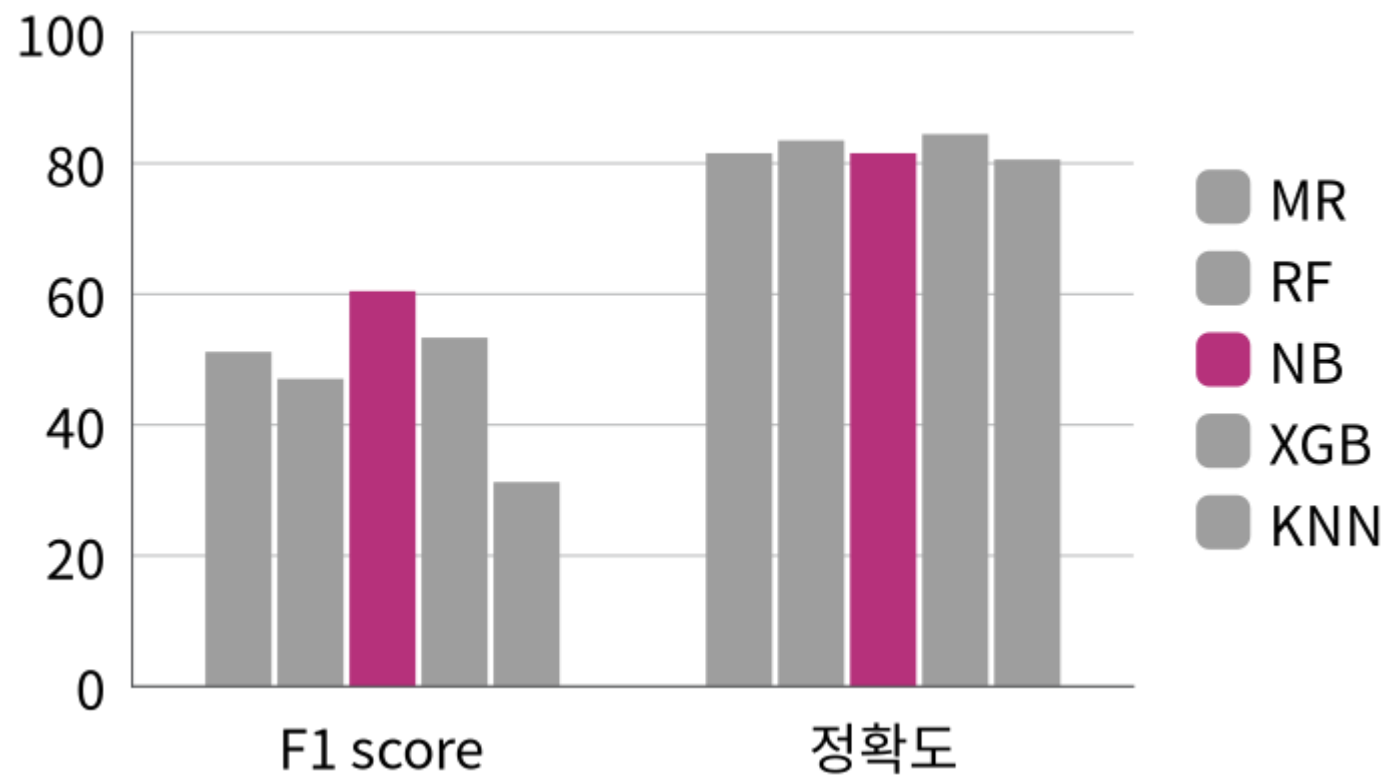
최종 모델 선정



일정 심각도

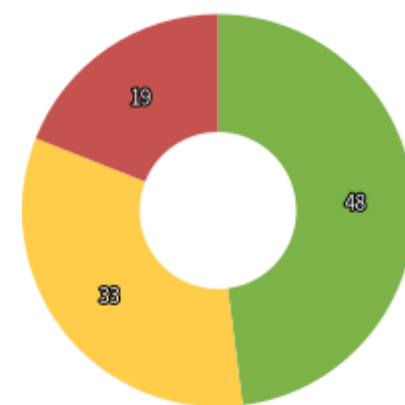
: 불균형 데이터

→ F1 score 기준



→ Naive Bayes

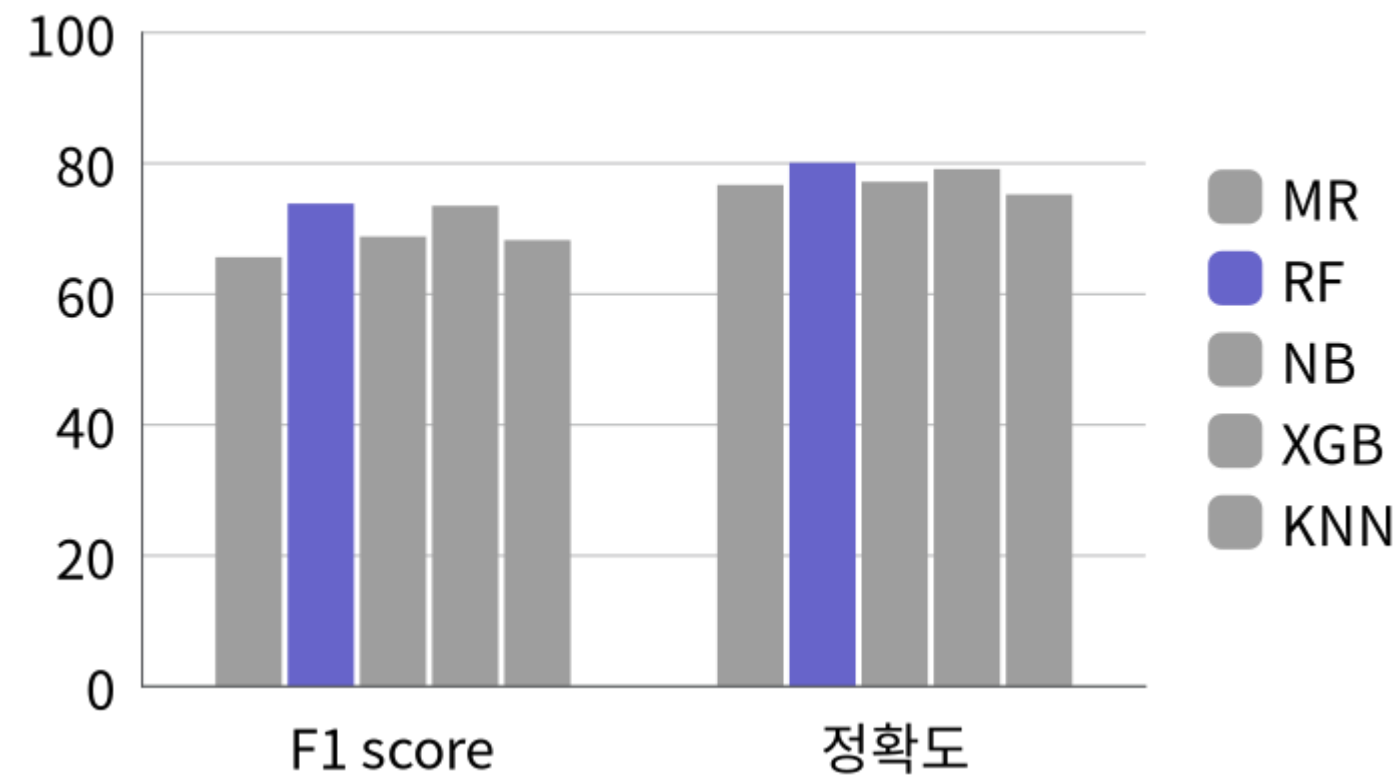
최종 모델 선정



금액 심각도

: 균형 데이터

→ 정확도 기준



→ Random Forest

R-Shiny 구현

http://127.0.0.1:4858 | Open in Browser

일정심각도 예측

프로젝트 분야를 선택하세요

화공 ▼

플랜트 종류를 선택하세요

육상 플랜트 ▼

Location을 선택하세요

SAUDI ▼

발주처를 선택하세요.

Saudi Aramco ▼

공종을 선택하세요.

배관 ▼

설계변경 사유를 입력하세요.

Non-metal 체결용 bolt 길이 변경(33 DV

실행

안전

http://127.0.0.1:4858 | Open in Browser

금액심각도 예측

프로젝트 분야를 선택하세요

화공 ▼

플랜트 종류를 선택하세요

육상 플랜트 ▼

Location을 선택하세요

SAUDI ▼

발주처를 선택하세요.

Saudi Aramco ▼

공종을 선택하세요.

배관 ▼

설계변경 사유를 입력하세요.

Non-metal 체결용 bolt 길이 변경(33 DV

실행

경계

경계	심각	안전
0.44	0.41	0.15

< Rshiny 일정심각도 >

< Rshiny 금액심각도 >

기대효과

새롭게 발생한 데이터

데이터입력
→

예측 모델

금액심각도 예측

프로젝트 분야를 선택하세요
항공

플랜트 종류를 선택하세요
육상 플랜트

Location을 선택하세요
DUBAI

발주처를 선택하세요
SABIC

공종을 선택하세요
배관

설계변경 사유를 입력하세요.
Pipe Rack 동쪽에 설치되는 Pump의 S

실행

안전

경계	심각	안전
0.38	0.05	0.57

정렬
→

활용

레코드번호	예측심각도
1	심각
2	심각
3	심각
...	...
40	경계
...
56	안전

sort
↓

경계	심각	안전
0.22	0.73	0.05

1. 심각한 사항 부터 처리 가능
2. 부당한 요구시 참고자료로 사용
3. 빠르고 합리적인 의사결정

프로젝트를 마치며



Q & A