

올바로시스템 노후 장비 교체 및 클라우드 인프라 증설

사용자(관리자) 매뉴얼



한국환경공단

Belltechsoft

<목차>

- 1 로그인(Log in)
- 2 모니터링
- 3 데이터 읽어 들이기
- 4 데이터 확인
- 5 데이터 전처리
- 6 모델 선택/학습
- 7 모델 평가/확정

1 로그인(Log In)

♣ 로그인(Log In)

① 로그인 : ID, Password 입력

Auto ML



Auto Machine Learning

Machine Learning과 Deep Learning을 기반으로 한 스마트한 AI Blue ML 솔루션

2022© Auto Machine Learning

1

LOGIN

Auto Machine Learning 로그인

☒ 아이디 저장

로그인

♣ 모니터링

- ① 전체 문서
- ② 데이터 전처리 완료
- ③ 모델 학습 완료
- ④ 모델 확정 완료



3

데이터 읽어 들이기

♣ 데이터 읽어 들이기

- ① [신규] 버튼 클릭 → ② <신규 파일 업로드> : 파일선택, 문서명
→ ③ [등록] 버튼 클릭
- ④ [선택] 버튼 클릭 → ⑤ 업로드한 [데이터 업로드 결과]의 상세 컬럼 내역이 표출됨
- ⑥ [삭제] 버튼 클릭 → ⑦ 알림창 '삭제 하시겠습니까?' ⑧ [예] 클릭 → '삭제되었습니다'

데이터 읽어 들이기

문서명

신규

등록

선택

삭제

데이터 업로드 결과

문서명	제출일자	등록일자	상태	세부
테스트 - 2022.11.21(01)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.02(01)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(11)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(09)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(08)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(06)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(05)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(04)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(03)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(02)	업로드	2022-11-18	선택	삭제

데이터 읽어 들이기

신규 파일 업로드

파일

문서명

등록

선택

삭제

데이터 업로드 결과

문서명	제출일자	등록일자	상태	세부
테스트 - 2022.11.21(01)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.02(01)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(11)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(09)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(08)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(06)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(05)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(04)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(03)	업로드	2022-11-18	선택	삭제
테스트 - 2022.11.01(02)	업로드	2022-11-18	선택	삭제

7



삭제 하시겠습니까?

예

아니요

8



삭제 되었습니다.

OK

4

데이터 확인

♣ 데이터 확인 (1/3)

- ① [조회] 버튼 클릭 → ② <문서 목록 데이터> 표출
 ③ [선택] 버튼 클릭 → ④ <컬럼별 데이터 분석 결과> 표출 → ⑤ [선택] 버튼 클릭
 ⑥ <데이터 분포(Boxplot)> 표출

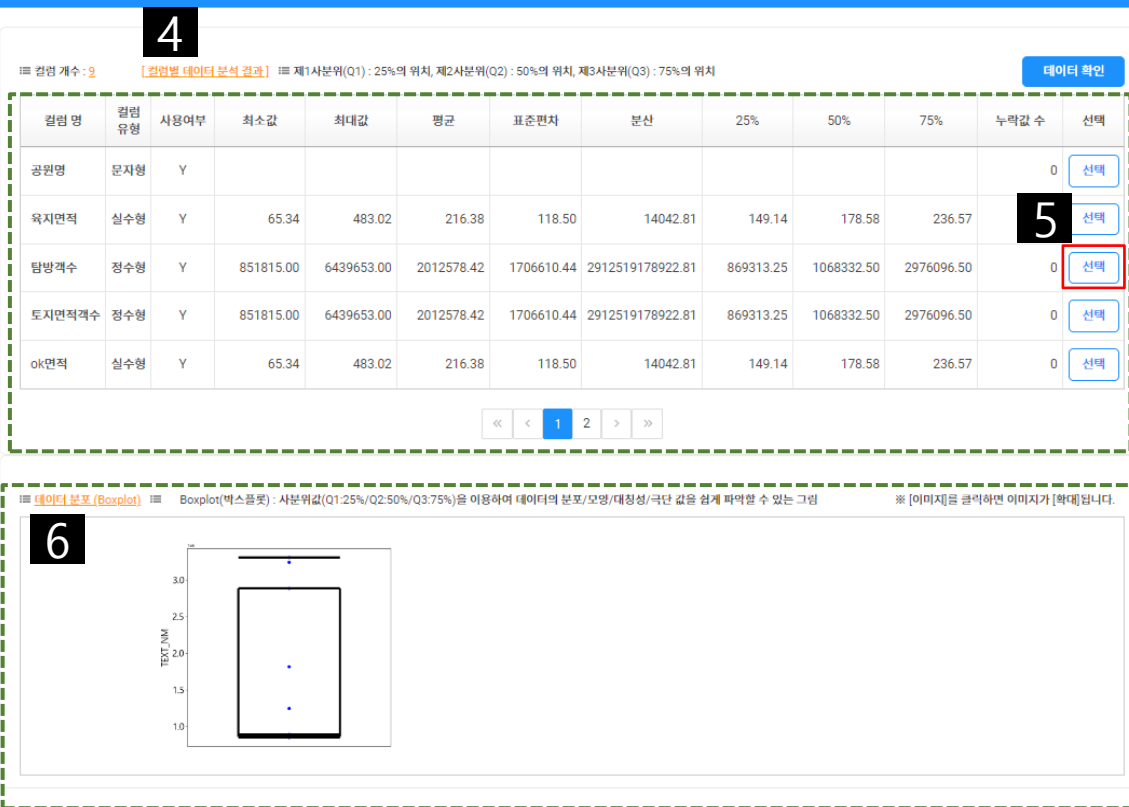
※ Boxplot(박스플롯) : 사분위값(Q1:25%/Q2:50%/Q3:75%)을 이용하여 데이터의 분포/모양/대칭성/극단 값을 쉽게 파악할 수 있는 그림

데이터 확인

문서명 처리상태

Total: 12

문서명	처리상태	등록일	선택
테스트 - 2022.11.21(01)	업로드	2022-11-21	선택
테스트 - 2022.11.02(01)	업로드	2022-11-18	선택
테스트 - 2022.11.01(11)	업로드	2022-11-18	선택
테스트 - 2022.11.01(10)	데이터 확인	2022-11-18	선택
테스트 - 2022.11.01(09)	업로드	2022-11-18	선택
테스트 - 2022.11.01(08)	업로드	2022-11-18	선택
테스트 - 2022.11.01(07)	데이터 확인	2022-11-18	선택
테스트 - 2022.11.01(06)	업로드	2022-11-18	선택
테스트 - 2022.11.01(05)	업로드	2022-11-18	선택
테스트 - 2022.11.01(04)	업로드	2022-11-18	선택



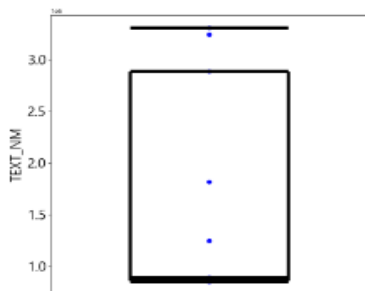
♣ 데이터 확인 (2/3)

⑦ <데이터 분포(Boxplot)>에서 표출된 이미지를 클릭 → ⑧ 확대된 [이미지]

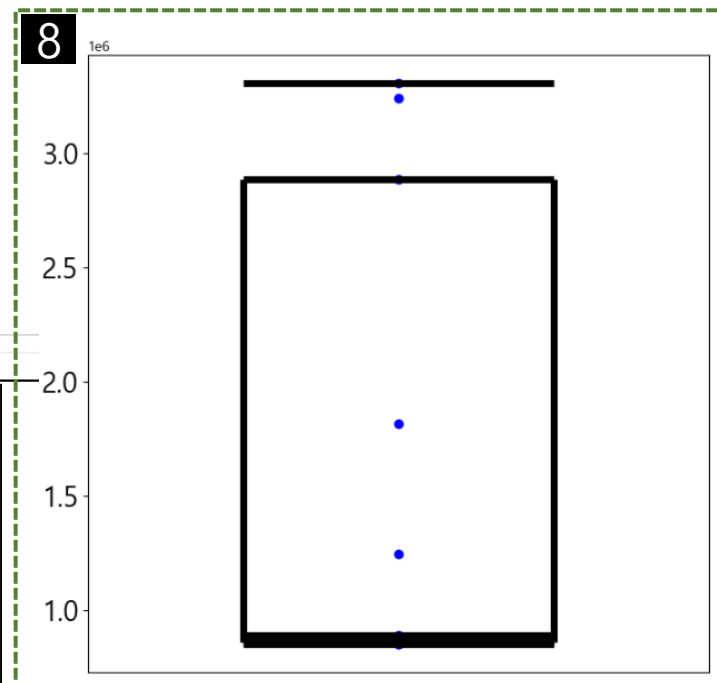
7

≡ 데이터 분포 (Boxplot) ≡ Boxplot(박스플롯) : 사분위값(Q1:25%/Q2:50%/Q3:75%)을 이용하여 데이터의 분포/모양/대칭성/극단 값을 쉽게 파악할 수 있는 그림

※ [이미지]를 클릭하면 이미지가 [확대]됩니다.



8



※ 참고사항

[사분위] 데이터를 4등분한 지점

- IQR(4분위수 범위)=Q3-Q1→ 데이터를 쌓아올렸을 때 25% 지점(1분위수)에 있는 데이터와 75% 지점(3분위수)에 있는 데이터의 차이

* 최소값 : 제 1사분위에서 1.5 IQR을 뺀 위치

* 최대값 : 제 3사분위에서 1.5 IQR을 더한 위치

* 제 1사분위(Q1) : 25%의 위치, 즉 전체 데이터 중 하위 25%에 해당하는 값

* 제 2사분위(Q2) : 50%의 위치로 중앙값(median)

* 제 3사분위(Q3) : 75%의 위치, 전체 데이터 중 상위 25%에 해당하는 값

4

데이터 확인

♣ 데이터 확인 (3/3)

- ⑨ [데이터 확인] 버튼 클릭 → ⑩ 알림창 '[데이터 확인] 처리 하시겠습니까?'
- ⑪ [예] 선택 → '[데이터 확인]이 정상적으로 처리되었습니다'

※ <문서 목록 데이터>의 처리상태가 '데이터 확인'으로 변경됨.

9

데이터 확인

== 컴럼 개수 : 9 [컴럼별 데이터 분석 결과] == 제1사분위(Q1) : 25%의 위치, 제2사분위(Q2) : 50%의 위치, 제3사분위(Q3) : 75%의 위치

컴럼 명	컴럼 유형	사용여부	최소값	최대값	평균	표준편차	분산	25%	50%	75%	누락값 수	선택
공관명	문자형	Y									0	선택
육자면적	실수형	Y	65.34	483.02	216.38	118.50	14042.81	149.14	178.58	236.57	0	선택
탕방객수	정수형	Y	851815.00	6439653.00	2012578.42	1706610.44	2912519178922.81	869313.25	1068332.50	2976096.50	0	선택
도자면적객수	정수형	Y	851815.00	6439653.00	2012578.42	1706610.44	2912519178922.81	869313.25	1068332.50	2976096.50	0	선택
ok면적	실수형	Y	65.34	483.02	216.38	118.50	14042.81	149.14	178.58	236.57	0	선택

<< < 1 2 > >>

10



[데이터 확인] 처리 하시겠습니까?

예

아니요

11



[데이터 확인]이 정상적으로 처리
되었습니다.

OK

♣ 데이터 전처리 (1/2)

- ① [조회] 버튼 클릭 → ② <문서 목록 데이터> 표출
- ③ [선택] 버튼 클릭 → ④ <컬럼별 데이터 전처리 결과> 표출
- ⑤ 목표변수로 사용하고자 하는 [목표변수] 선택란에 check ※ [목표변수는 반드시 1 컬럼만 선택해야만 모델학습이 가능]
- ⑥ [목표변수 적용] 버튼 클릭 : 알림창 '목표변수 적용' 처리를 하시겠습니까?
※ <문서 목록 데이터>의 처리상태가 '데이터 전처리'로 변경됨.

데이터 전처리

문서명 처리상태

Total: 32 1

문서명	처리상태	등록일	선택
테스트 - 2022.11.21(01)	업로드	2022-11-21	<input type="button" value="선택"/>
테스트 - 2022.11.02(01)	업로드	2022-11-18	<input type="button" value="선택"/>
테스트 - 2022.11.01(11)	업로드	2022-11-18	<input type="button" value="선택"/>
테스트 - 2022.11.01(10)	데이터 확인	2022-11-18	<input type="button" value="선택"/>
테스트 - 2022.11.01(09)	업로드	2022-11-18	<input type="button" value="선택"/>
테스트 - 2022.11.01(08)	데이터 확인	2022-11-18	<input type="button" value="선택"/>
테스트 - 2022.11.01(07)	데이터 확인	2022-11-18	<input type="button" value="선택"/>
테스트 - 2022.11.01(06)	업로드	2022-11-18	<input type="button" value="선택"/>
테스트 - 2022.11.01(05)	업로드	2022-11-18	<input type="button" value="선택"/>
테스트 - 2022.11.01(04)	업로드	2022-11-18	<input type="button" value="선택"/>

« < 1 2 3 4 > »

6 목표변수 적용 ※ [목표변수는 반드시 1 컬럼만 선택해야 모델학습이 가능합니다] ※ 목표변수(결과변수) : 추정하거나 예측하고 싶은 목적 데이터 (예) 등급/가점/성별/학력

컬럼명	컬럼유형	최소값	최대값	평균	표준편차	분산	25%	50%	75%	누락값 수	목표변수 선택
ok수정구	실수형	365.34	9173.53	4241.38	2942.37	8657568.40	1973.17	3675.51	6494.41	0	<input type="checkbox"/>
수지구 면적	실수형	65.34	483.02	216.38	118.50	14042.81	149.14	178.58	236.57	0	<input type="checkbox"/>
송파 면적	실수형	65.34	483.02	216.38	118.50	14042.81	149.14	178.58	236.57	0	<input type="checkbox"/>
y	문자형									0	<input checked="" type="checkbox"/>

4 5 « < 1 2 > »

7 데이터 정제 (Data Cleansing) 데이터 전처리 적용 ※ [결측값, Missing Value] 알려지지 않고, 수집되지 않거나 잘못 입력된 데이터 세트의 값
※ [이상값, 극단값, Outlier] 특정 데이터 변수의 분포에서 비정상적으로 벗어난 값

[데이터 전처리]구분	<input type="checkbox"/> 제거 여부 선택	<input type="checkbox"/> 사용 여부 선택
중복값 처리	<input type="checkbox"/>	<input checked="" type="checkbox"/>
결측값 처리	<input type="checkbox"/>	<input checked="" type="checkbox"/>
이상값 처리	<input type="checkbox"/>	<input checked="" type="checkbox"/>

♣ 데이터 전처리 (2/2)

- ⑦ 전처리 적용 여부 선택 -> 데이터 전처리하고자 하는 각각 구분에 따라 선택 란에 check
 ⑧ 사용 여부 선택 -> 체크가 된 경우 : 사용, 체크가 안된 경우 : 미사용
 ⑨ [데이터 전처리 적용] 버튼 클릭 : 알림창 '[데이터 전처리] 작업을 하시겠습니까?'
 → [데이터 전처리] 작업이 정상적으로 처리 되었습니다.

≡ 데이터 정제 (Data Cleansing)

9

데이터 전처리 적용

※ [결측값, Missing Value] 알려지지 않고, 수집되지 않거나 잘못 입력된 데이터 세트의 값

※ [이상값, 극단값, Outlier] 특정 데이터 변수의 분포에서 비정상적으로 벗어난 값

[데이터 전처리]구분	7	8
	□ 전처리 적용 여부 선택	□ 사용 여부 선택
중복값 처리	<input type="checkbox"/>	<input checked="" type="checkbox"/>
결측값 처리	<input type="checkbox"/>	<input checked="" type="checkbox"/>
이상값 처리	<input type="checkbox"/>	<input checked="" type="checkbox"/>

※ 데이터 전처리 적용

- 1) 중복값 처리 : [처리방법] 중복되는 값 중 첫번째 값만 제외하고 나머지를 제거
- 2) 결측값 처리 : Missing Value - 알려지지 않고, 수집되지 않거나 잘못 입력된 데이터 세트의 값 [처리방법] 평균(mean)값
- 3) 이상값 처리 : Outlier - 이상값 = 극단값 = 이상점 : 특정 데이터 변수의 분포에서 비정상적으로 벗어난 값 [처리방법] 평균(mean)값

♣ 모델 선택/학습 (1/2)

- ① [조회] 버튼 클릭 → ② <문서 목록 데이터> 표출
- ③ 학습할 문서에서 [선택] 버튼 클릭 → ④ <모델 학습 목록> 표출
- ⑤ 분류(Classification) / 군집(Clustering) 모델 중에서 학습하고자 하는 [모델] 클릭
→ 알림창 '[분류 모델]을 학습하시겠습니까?' → [예] 클릭 → '[분류 모델] 학습이 정상적으로 처리 되었습니다.'

모델 선택/학습

문서명 처리상태

Total: 21 조회

문서명	처리상태	등록일	선택
테스트 - 2022.11.01(10)	데이터 전처리	2022-11-18	선택
테스트 - 2022.11.17 (15개)	모델 학습	2022-11-17	선택
테스트 - 2022.11.17 (14개)	모델 학습	2022-11-17	선택
테스트 - 2022.11.17 (13개)	모델 학습	2022-11-17	선택
테스트 - 2022.11.17 (12개)	모델 학습	2022-11-17	선택
테스트 - 2022.11.16(04)	모델 학습	2022-11-16	선택
테스트 - 2022.11.16(05)	모델 학습	2022-11-16	선택
테스트 - 2022.11.16(02)	데이터 전처리	2022-11-16	선택
테스트 - 2022.11.16(01)	데이터 전처리	2022-11-16	선택
테스트 - 2022.11.15(01)	모델 학습	2022-11-15	선택

« < 1 2 3 > »

분류(Classification) 모델 그레디언트 부스팅(Gradient Boost) 의사 결정 트리 Random Forest SVM(Support Vector Machine) 신경망(Neural Network)

군집(Clustering) 모델 K-평균 군집화(K-means Clustering)

※ 아래 화면의 [이미지]를 클릭하면 이미지가 [확대]됩니다.

Total: 5 모델 학습 목록

※ 정밀도(Precision) : 모델이 True라고 분류한 것 중에서 실제 True인 것의 비율
 ※ 민감도(Sensitivity) : 실제 True인 것 중에서 모델이 True라고 예측한 것의 비율
 ※ F1-score : 정밀도와 민감도의 조화평균 -> 2*(정밀도*민감도)/(정밀도+민감도)

등록된 정보가 없습니다.

Confusion Matrix (혼동행렬) / (신경망) 훈련 데이터 대 검증 데이터 손실 그래프

※ (혼동/오차행렬) 학습을 통한 예측성능을 측정하기 위해 예측 값과 실제 값을 비교하기 위한 표

ROC Curve ※ 특이도(Specificity) : 실제 진단결과가 음성 중에 음성을 음성이라고 맞춘 비율
 ※ (수신자 판단 곡선) 모델의 효율성을 민감도와 특이도를 이용하여 그래프로 나타낸 것

♣ 모델 선택/학습 (2/2)

- ⑥ 알림창 '[분류 모델] 학습이 정상적으로 처리 되었습니다.' -> <모델 학습 목록>에 학습된 모델 표출
(예) 그레디언트 부스팅(Gradient Boost)을 선택한 경우
- ⑦ [선택] 버튼 클릭 → ⑧ Confusion Matrix / 훈련 데이터 대 검증 데이터 손실 그래프
⑨ Roc Curve

6

Total: 1

모델 학습 목록

※정밀도(Precision) : 모델이 True라고 분류한 것 중에서 실제 True인 것의 비율

※ 민감도(Sensitivity) : 실제 True인 것 중에서 모델이 True라고 예측한 것의 비율

※ F1-score : 정밀도와 민감도의 조화평균 -> $2 * (\text{정밀도} * \text{민감도}) / (\text{정밀도} + \text{민감도})$

학습 구분	정밀도	민감도	F1-score	학습일자	학습횟수	선택
Gradient Boost	0.5	0.5	0.5	2022-11-25	1	선택

7

※ 참고사항

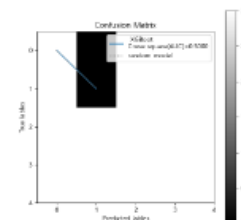
- (1) 정밀도(Precision) : 모델이 True라고 분류한 것 중에서 실제 True인 것의 비율
(2) 민감도(Sensitivity) : 실제 True인 것 중에서 모델이 True라고 예측한 것의 비율
(3) F1-score : 정밀도와 민감도의 조화평균
-> $2 * (\text{정밀도} * \text{민감도}) / (\text{정밀도} + \text{민감도})$

8

Confusion Matrix (혼동행렬)

/ (신경망) 훈련 데이터 대 검증 데이터 손실 그래프

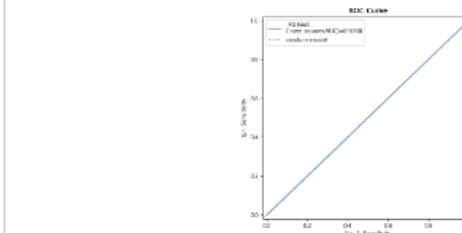
※(혼동/오차행렬) 학습을 통한 예측성능을 측정하기 위해 예측 값과 실제 값을 비교하기 위한 표



ROC Curve

※ 특이도(Specificity) : 실제 진단결과가 음성 중에 음성을 음성이라고 맞춘 비율

※ (수신자 판단 곡선) 모델의 효율성을 민감도와 특이도를 이용하여 그래프로 나타낸 것



♣ 모델 평가/확정 (1/2)

- ① [조회] 버튼 클릭 → ② <문서 목록 데이터> 표출
- ③ <문서 목록 데이터> 목록에서 [선택] 버튼 클릭 → ④ <모델 학습 목록> 표출
- ⑤ <모델 학습 목록>에서 [선택] 버튼 클릭 → ⑥ Confusion Matrix / Roc Curve

모델 평가/확정

문서 명

처리상태

Total : 23

1

조회

2

문서 명	처리상태	등록일	선택
테스트 - 2022.11.01(10)	모델 학습	2022-11-18	선택
테스트 - 2022.11.17 (15개)	모델 학습	2022-11-17	선택
테스트 - 2022.11.17 (14개)	모델 학습	2022-11-17	선택
테스트 - 2022.11.17 (13개)	모델 학습	2022-11-17	선택
테스트 - 2022.11.17 (12개)	모델 학습	2022-11-17	선택
테스트 - 2022.11.17 (11개)	모델 학습	2022-11-17	선택
테스트 - 2022.11.16(04)	모델 학습	2022-11-16	선택
테스트 - 2022.11.16(05)	모델 학습	2022-11-16	선택
테스트 - 2022.11.15(01)	모델 학습	2022-11-15	선택
테스트 - 2022.11.11(03)	모델 학습	2022-11-11	선택

<< < 1 2 3 > >>

3

4

모델 학습 목록

※정밀도(Precision) : 모델이 True라고 분류한 것 중에서 실제 True인 것의 비율
 ※ 민감도(Sensitivity) : 실제 True인 것 중에서 모델이 True라고 예측한 것의 비율
 ※ F1-score : 정밀도와 민감도의 조화평균 → $2 * (\text{정밀도} * \text{민감도}) / (\text{정밀도} + \text{민감도})$

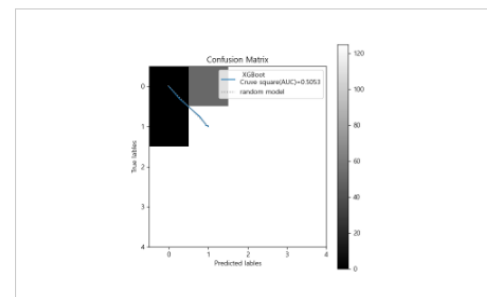
학습 구분	정밀도	민감도	F1-score	학습일자	학습횟수	확정일자	선택	확정
Gradient Boost	0.71	0.71	0.71	2022-11-17	2		선택	확정
의사결정트리	0.71	0.71	0.71	2022-11-11	1		선택	확정
Random Forest	0.71	0.71	0.71	2022-11-11	1		선택	확정
SVM	0.7	0.7	0.7	2022-11-11	1		선택	확정
신경망	0	0	0	2022-11-17	1		선택	확정

5

6

Confusion Matrix (혼동행렬) / (신경망) 훈련 데이터 대 검증 데이터 손실 그래프

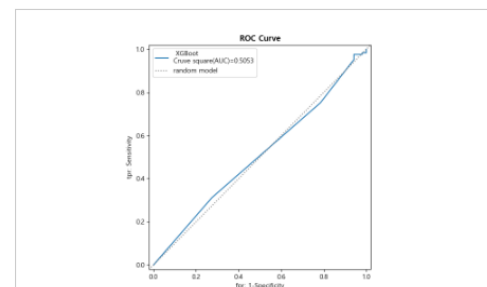
※오차행렬 : 학습을 통한 예측성능을 측정하기 위해 예측결과 값과 실제 값을 비교하기 위한 표



ROC Curve

※ [이미지]를 클릭하면 이미지가 확대됩니다.

※ (수신자 판단 곡선) 모델의 효율성을 민감도와 특이도를 이용하여 그래프로 나타낸 것.



♣ 모델 평가/확정 (2/2)

- ⑦ <모델 학습 목록>에서 [확정]하고자 하는 모델 확정 처리 (예) Random Forest를 확정처리
 → [확정] 버튼 클릭 → 알림창 ‘[모델 확정] 처리 하시겠습니까?’ → [예] 클릭 → ‘[모델 확정]이 정상적으로 처리 되었습니다.’
- ⑧ <모델 학습 목록>에 [확정] 처리된 내역 표출
- ⑨ <문서 목록 데이터>의 처리상태가 ‘모델 확정’으로 변경됨.

이 Total: 5 모델 학습 목록

※ 정밀도(Precision): 모델이 True라고 분류한 것 중에서 실제 True인 것의 비율
 ※ 민감도(Sensitivity): 실제 True인 것 중에서 모델이 True라고 예측한 것의 비율
 ※ F1-score: 정밀도와 민감도의 조화평균 → $2 * (\text{정밀도} * \text{민감도}) / (\text{정밀도} + \text{민감도})$

학습 구분	정밀도	민감도	F1-score	학습일자	학습횟수	확정일자	선택	확정
Gradient Boost	0.71	0.71	0.71	2022-11-17	2		<input type="button" value="선택"/>	<input type="button" value="확정"/>
의사결정트리	0.71	0.71	0.71	2022-11-11	1		<input type="button" value="선택"/>	<input type="button" value="확정"/>
Random Forest	0.71	0.71	0.71	2022-11-11	1		<input type="button" value="선택"/>	<input type="button" value="확정"/>
SVM	0.7	0.7	0.7	2022-11-11	1		<input type="button" value="선택"/>	<input type="button" value="확정"/>
신경망	0	0	0	2022-11-17	1		<input type="button" value="선택"/>	<input type="button" value="확정"/>



학습 구분	정밀도	민감도	F1-score	학습일자	학습횟수	확정일자	선택	확정
Gradient Boost	0.71	0.71	0.71	2022-11-17	2		<input type="button" value="선택"/>	<input type="button" value="확정"/>
의사결정트리	0.71	0.71	0.71	2022-11-11	1		<input type="button" value="선택"/>	<input type="button" value="확정"/>
Random Forest	0.71	0.71	0.71	2022-11-11	1	2022-11-25	<input type="button" value="선택"/>	<input type="button" value="확정"/>
SVM	0.7	0.7	0.7	2022-11-11	1		<input type="button" value="선택"/>	<input type="button" value="확정"/>
신경망	0	0	0	2022-11-17	1		<input type="button" value="선택"/>	<input type="button" value="확정"/>

이 Total: 23

조회

문서 명	처리상태	등록일	선택
테스트 - 2022.11.01(10)	모델 학습	2022-11-18	<input type="button" value="선택"/>
테스트 - 2022.11.17 (15개)	모델 학습	2022-11-17	<input type="button" value="선택"/>
테스트 - 2022.11.17 (14개)	모델 학습	2022-11-17	<input type="button" value="선택"/>
테스트 - 2022.11.17 (13개)	모델 학습	2022-11-17	<input type="button" value="선택"/>
테스트 - 2022.11.17 (12개)	모델 학습	2022-11-17	<input type="button" value="선택"/>
테스트 - 2022.11.17 (11개)	모델 확정	2022-11-17	<input type="button" value="선택"/>
테스트 - 2022.11.16(04)	모델 학습	2022-11-16	<input type="button" value="선택"/>
테스트 - 2022.11.16(05)	모델 학습	2022-11-16	<input type="button" value="선택"/>
테스트 - 2022.11.15(01)	모델 학습	2022-11-15	<input type="button" value="선택"/>
테스트 - 2022.11.11(03)	모델 확정	2022-11-11	<input type="button" value="선택"/>