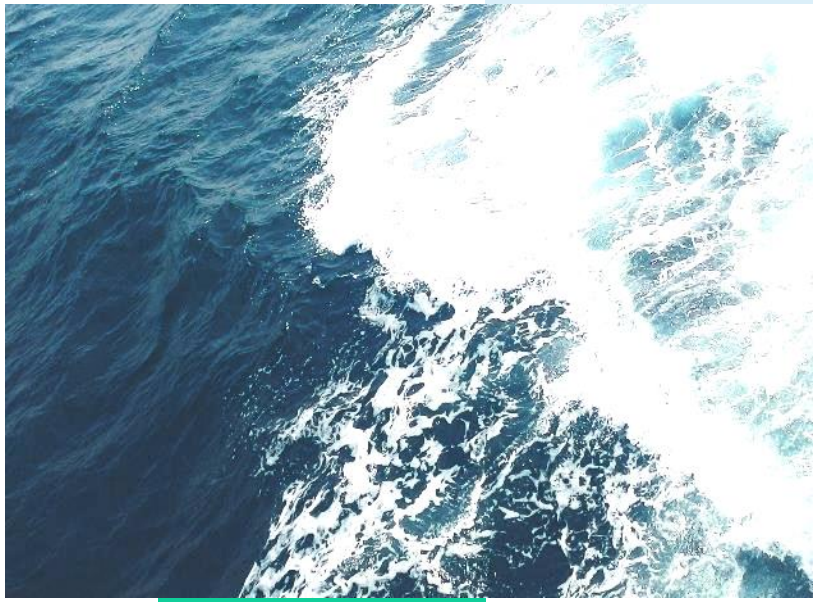
Section2 Project

적조현상 발생여부 예측



AI_16_한승희

Table of contents

01

분석 배경

- 필요성
- 분석목적 및 문제정의

필요성

- **적조(red tide) 정의** : 유해적조 생물의 일시적인 대량 번식으로 바다물의 색이 적색, 황색, 적갈색으로 변하는 현상

- **적조발생으로 인한 피해**

  연안 환경 및 생태계에 악영향을 미치며, 양식장의 어패류를 집단 폐사시켜 수산업에 큰 피해를 일으킴

적조 피해금액이 점차 낮아지는 추세이지만 여전히 억원 단위에 달함.

또한 피해가 없더라도 매년 방재 및 예방에 많은 예산이 소모되고 있음.

참고 4 연도별 적조 발생 현황

연도	최초 발생일	최초 발생지역	발생범위	지속일 (일)	최대밀도 (개체수/mL)	피해액 (억원)
'95	8.29	고흥	완도~강릉	54	30,000	764
'96	9.5	고흥, 여천	완도~기장	28	23,000	21
'97	8.25	고흥	완도~울진	29	20,000	15
'98	8.30	고흥	완도~거제	34	20,000	16
'99	8.11	고흥	완도~울진	54	43,000	3.2
'00	8.22	여수, 남해	고흥~기장	29	15,000	2.6
'01	8.14	여수	완도~삼척	42	32,000	84
'02	8.2	여수	완도~울진	55	30,000	49
'03	8.13	여수~남해	진도~강릉	62	48,000	215
'04	8.5	거제	완도~거제	30	5,800	1.2
'05	7.19	고흥	완도~거제	58	25,000	10.6
'06	8.6	여수	완도~남해	37	33,500	0.7
'07	7.31	고흥	완도~울진	50	32,500	115
'08	7.30	고흥	완도~울산	62	7,300	-
'09	10.28	여수	여수~통영	20	1,660	-
'10	9.17	통영	통영	3	1,300	-
'11	적조 미 발생					
'12	7.27	고흥	완도~거제, 태안	75	23,000	44
'13	7.17	여수, 통영	고흥~양양	51	34,800	247
'14	7.24	경남 고성	완도~삼척	86	20,000 (포항9.13, 9.19)	74
'15	8.2	경남 통영	진도~울진	56	32,000	53
'16	8.16	전남 여수(코) 고흥~완도(카)	여수~완도	14	2,200(코) 1,280(카)	43
'17	적조 미 발생					
'18	7.23	전남 여수~ 경남 남해	고흥~거제	28	4,500	2.7

* '16년 : (코) 코클로디니움 적조, (카) 카레니아 적조 / '16년 카레니아 적조로 전복 피해(전남)
* '18년 : 2건의 피해발생(어류 178천미, 274백만원) 하였으나 보험 처리.

https://www.busan.com/view/busan/view.php?code=20130730000147

2019년 적조대응 종합대책(해양수산부)

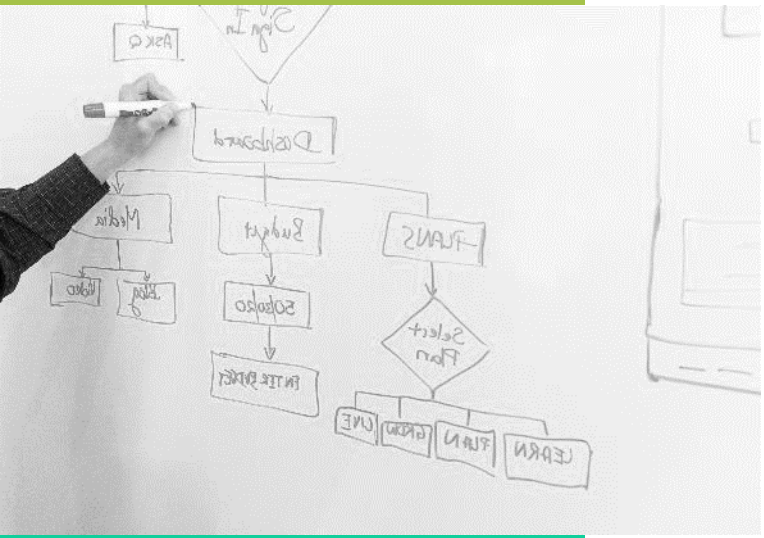
분석목적 및 문제정의

❖ 분석목적

- 기후온난화로 인한 수온상승, 기후변화, 해양환경 변화에 대한 적조 발생여부 예측 모델을 마련하고자 함

❖ 문제정의

- 적조발생에 영향을 미치는 기상 요인, 해양 요인 및 그 관계성 파악
- 적조가 발생할 가능성이 큰 환경조건의 조합 파악
- 기상요인과 해양요인 데이터를 바탕으로 적조가 발생할 지 여부를 예측



02

분석 개요

- 분석절차, 활용모델
- 가설설정

분석절차, 활용모델

데이터 수집

적조발생예측데이터

Date	관측일자
Temp	수온(℃)
Rainfall_day	일강수량(mm)
WindSpeed	평균 풍속(m/s)
Atm	기압(hPa)
DO	용존산소량(㎎/ℓ)
COD	화학산소요구량(㎎/ℓ)
pH	pH
TN	총질소(㎎/ℓ)
TP	총인(㎎/ℓ)
Min_Density	적조 최저밀도(개체/㎖)
Max_Density	적조 최고밀도(개체/㎖)

환경빅데이터플랫폼, 국립수산과학원

+

기상청

ASOS_여수

Date	관측일자
Temp	평균기온(℃)
Rainfall_Dur	강수 계속시간(hr)
Rainfall	일강수량(mm)
Wind_Speed	평균 풍속(m/s)
Sunshine	합계 일조시간(hr)

국가통계포털

하천수의_수질현황영산강권역

Date(yy-mm)	관측일자
DO	용존산소 DO (㎎/ℓ)
COD	화학적산소요구량 COD (㎎/ℓ)
pH	수소이온농도 (pH)
TN	총질소 T-N (㎎/ℓ)
TP	총인 T-P (㎎/ℓ)

전처리

EDA

특성공학

최종 DATA

Cochlo_yn	적조발생여부
m_Temp	직전7일평균 수온
m_Rainfall_Dur	직전7일평균 강수 계속시간
m_Rainfall_day	직전7일평균 강수량
m_WindSpeed	직전7일평균 풍속
m_Atm	직전7일평균 기압
m_Sunshine	직전7일평균 일조시간
m_DO	직전7일평균 용존산소량
m_COD	직전7일평균 화학적산소요구량
m_pH	직전7일평균 수소이온농도
m_TN	직전7일평균 총질소
m_TP	직전7일평균 총인

모형 구축 및 검증

- 로지스틱 회귀분석
- 랜덤포레스트

가설설정

1. 여름철에 적조현상이 빈번하게 발생한다.

   여름철 → 수온↑, 일조시간↑, 강수량↑

2. 이 밖에 해양환경에서 적조 발생에 기여하는 요인이 있다.

03

활용 데이터

- 데이터 정보
- 전처리
- 특성공학

데이터 정보

❖ data | main data

	Date	Temp	Rainfall_day	WindSpeed	Atm	Solidity	DO	COD	pH	Turbidity	TN	TP	Min_Density	Max_Density	Cochlo_YN
0	2014-01-01	7.700	0.000	5.263	1,015.147	29.287	10.071	1.168	7.843	262.742	0.290	0.026	0	0	0
1	2014-01-02	7.700	0.000	1.219	1,019.767	29.270	10.001	1.297	7.859	194.517	0.290	0.026	0	0	0
2	2014-01-03	7.700	0.000	1.238	1,016.513	29.274	9.984	1.163	7.846	86.668	0.290	0.026	0	0	0
3	2014-01-04	7.800	0.000	2.419	1,016.210	29.235	9.946	0.845	7.850	68.625	0.290	0.026	0	0	0
4	2014-01-05	7.700	0.000	2.688	1,020.507	29.284	9.889	0.818	7.869	85.875	0.290	0.028	0	0	0

❖ data_asos | 기상관측 자료

	station_num	station	Date	Temp	Rainfall_Dur	Rainfall	Wind_Speed	Sunshine
0	168	여수	2014-01-01	7.200	NaN	NaN	6.400	8.500
1	168	여수	2014-01-02	6.400	NaN	NaN	1.500	7.100
2	168	여수	2014-01-03	6.700	NaN	NaN	1.700	7.300
3	168	여수	2014-01-04	6.000	NaN	NaN	4.300	8.900
4	168	여수	2014-01-05	4.500	NaN	NaN	3.500	8.700

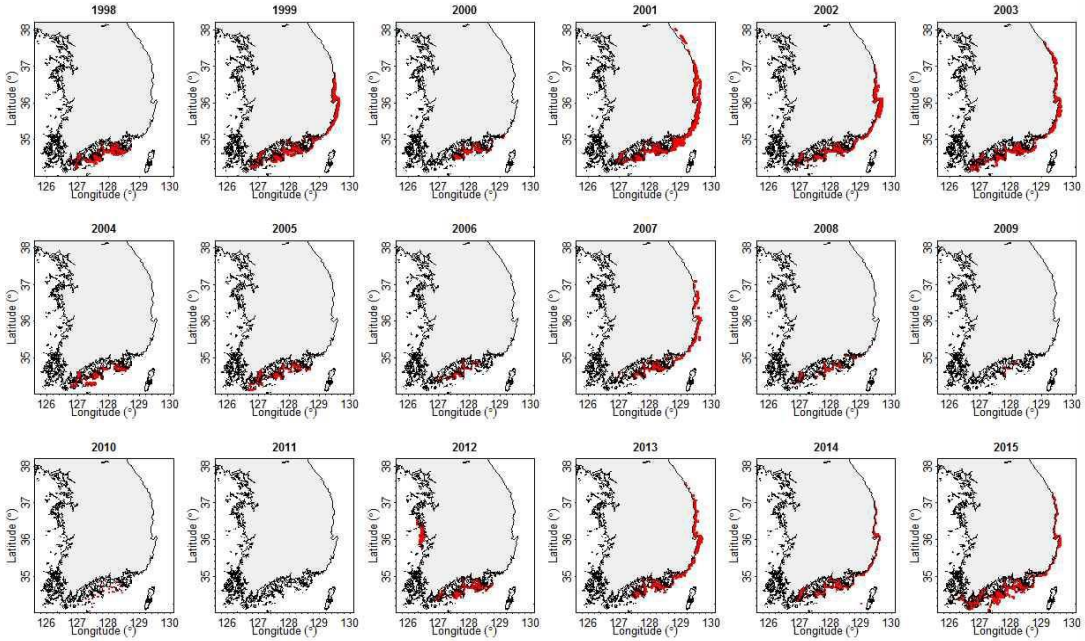
- 수집기간 : 2014-01-01 ~ 2019-08-31
- data size : 2068

❖ data_water_sup | 해양환경 자료 (데이터 보완용)

	시점	수계별 (1)	수계별 (2)	수계별 (3)	용존산소 DO (mg/ℓ)	화학적산소요구량 COD (mg/ℓ)	수온 (℃)	수소이온농도 (pH)	총질소 T-N (mg/ℓ)	총인 T-P (mg/ℓ)	TOC (mg/ℓ)
0	2014-05-01	기타	금산면	오천천	9.500	1.900	17.400	8.700	1.335	0.007	1.1
1	2014-05-01	기타	이사천	이사천1	11.200	2.800	18.000	8.400	1.041	0.026	1.1
2	2014-05-01	기타	이사천	이사천2	9.500	2.100	16.200	7.900	1.120	0.018	1.3
3	2014-05-01	기타	수어천	수어천	8.800	2.100	21.000	7.400	1.651	0.015	0.8
4	2014-05-01	기타	수어천	수어천1	8.400	0.600	21.000	7.500	0.922	0.030	1.8

※ data 셋에서 장기간 동일값으로 나온 구간이 존재해 이를 보완하기 위함

데이터 선택 배경



위성빅데이타기반적조탐지및발생경향분석연구(2015,한국해양과학기술원)

적조정보시스템

- 적조발생은 매년 남해안에서 시작되는 것으로 조사됨(남해에서 발생된 적조가 대마 난류에 의해 동해로 확산)
- 적조생물밀도가 클수록 피해가 크다고 확정할 수 없고, 적조원인생물이 유해종류인지에 따라 피해가 달라짐 → Cochlodinium

전처리 - 데이터셋 취합

data | main data

	Date	Temp	Rainfall_day	WindSpeed	Atm	Solidity	DO	COD	pH	Turbidity	TN	TP	Min_Density	Max_Density	Cochlo_YN
0	2014-01-01	7.700	0.000	5.263	1,015.147	29.287	10.071	1.168	7.843	262.742	0.290	0.026	0	0	0
1	2014-01-02	7.700	0.000	1.219	1,019.767	29.270	10.001	1.297	7.859	194.517	0.290	0.026	0	0	0
2	2014-01-03	7.700	0.000	1.238	1,016.513	29.274	9.984	1.163	7.846	86.668	0.290	0.026	0	0	0
3	2014-01-04	7.800	0.000	2.419	1,016.210	29.235	9.946	0.845	7.850	68.625	0.290	0.026	0	0	0
4	2014-01-05	7.700	0.000	2.688	1,020.507	29.284	9.889	0.818	7.869	85.875	0.290	0.028	0	0	0

data_asos | 기상관측 자료

	station_num	station	Date	Temp	Rainfall_Dur	Rainfall	Wind_Speed	Sunshine
0	168	여수	2014-01-01	7.200	NaN	NaN	6.400	8.500
1	168	여수	2014-01-02	6.400	NaN	NaN	1.500	7.100
2	168	여수	2014-01-03	6.700	NaN	NaN	1.700	7.300
3	168	여수	2014-01-04	6.000	NaN	NaN	4.300	8.900
4	168	여수	2014-01-05	4.500	NaN	NaN	3.500	8.700

data_water_sup | 해양환경 자료 (데이터 보완용)

	시점	수계별 (1)	수계별 (2)	수계별 (3)	용존산소 DO (mg/ℓ)	화학적산소요구량 COD (mg/ℓ)	수온 (℃)	수소이온농도 (pH)	총질소 T-N (mg/ℓ)	총인 T-P (mg/ℓ)	TOC (mg/ℓ)
0	2014-05-01	기타	금산면	오천천	9.500	1.900	17.400	8.700	1.335	0.007	1.1
1	2014-05-01	기타	이사천	이사천1	11.200	2.800	18.000	8.400	1.041	0.026	1.1
2	2014-05-01	기타	이사천	이사천2	9.500	2.100	16.200	7.900	1.120	0.018	1.3
3	2014-05-01	기타	수어천	수어천	8.800	2.100	21.000	7.400	1.651	0.015	0.8
4	2014-05-01	기타	수어천	수어천1	8.400	0.600	21.000	7.500	0.922	0.030	1.8

- 데이터셋간 중복되는 컬럼, 자료가 부족한 컬럼, 예측에 필요하지 않은 컬럼 제거
- 날짜 컬럼을 키값으로 하여 데이터 취합 → 컬럼 순서 정리

전처리 - 데이터셋 취합결과

data

	Date	Temp	Rainfall_Dur	Rainfall_day	WindSpeed	Atm	Sunshine	DO	COD	pH	TN	TP	Min_Density	Max_Density
2048	2019-08-12	26.500	5.780	2.100	3.299	1,001.549	0.100	5.029	4.479	7.996	0.136	0.040	0	0
2049	2019-08-13	26.700	0.000	0.000	1.153	1,001.720	11.500	5.681	4.416	7.876	0.097	0.036	0	0
2050	2019-08-14	27.200	2.920	0.000	3.682	997.747	11.300	6.270	1.786	8.140	0.112	0.045	0	0
2051	2019-08-15	27.500	2.420	0.100	4.052	990.110	4.800	6.159	5.138	8.244	0.1			0
2052	2019-08-16	26.900	0.000	0.000	5.205	995.604	10.400	5.628	6.159	8.165	0.1			
2053	2019-08-17	25.900	0.000	0.000	3.156	1,002.491	12.400	5.661	6.189	8.129	0.2			
2054	2019-08-18	25.400	0.000	0.000	1.801	1,005.505	10.800	5.562	6.050	8.075	0.1			
2055	2019-08-19	25.300	0.000	0.000	1.520	1,008.088	7.200	5.346	6.292	8.071	0.1			
2056	2019-08-20	25.400	0.000	0.000	2.081	1,008.771	4.500	5.041	5.518	8.029	0.1			
2057	2019-08-21	25.400	6.700	18.400	1.186	1,008.447	0.600	4.530	6.000	7.981	0.1			
2058	2019-08-22	25.500	7.130	13.500	4.691	1,005.954	4.200	4.789	4.677	8.028	0.1			
2059	2019-08-23	25.300	0.680	0.000	3.212	1,008.278	2.900	4.951	5.957	8.053	0.1			
2060	2019-08-24	25.600	0.000	0.000	2.344	1,010.355	5.000	4.664	5.947	8.017	0.1			

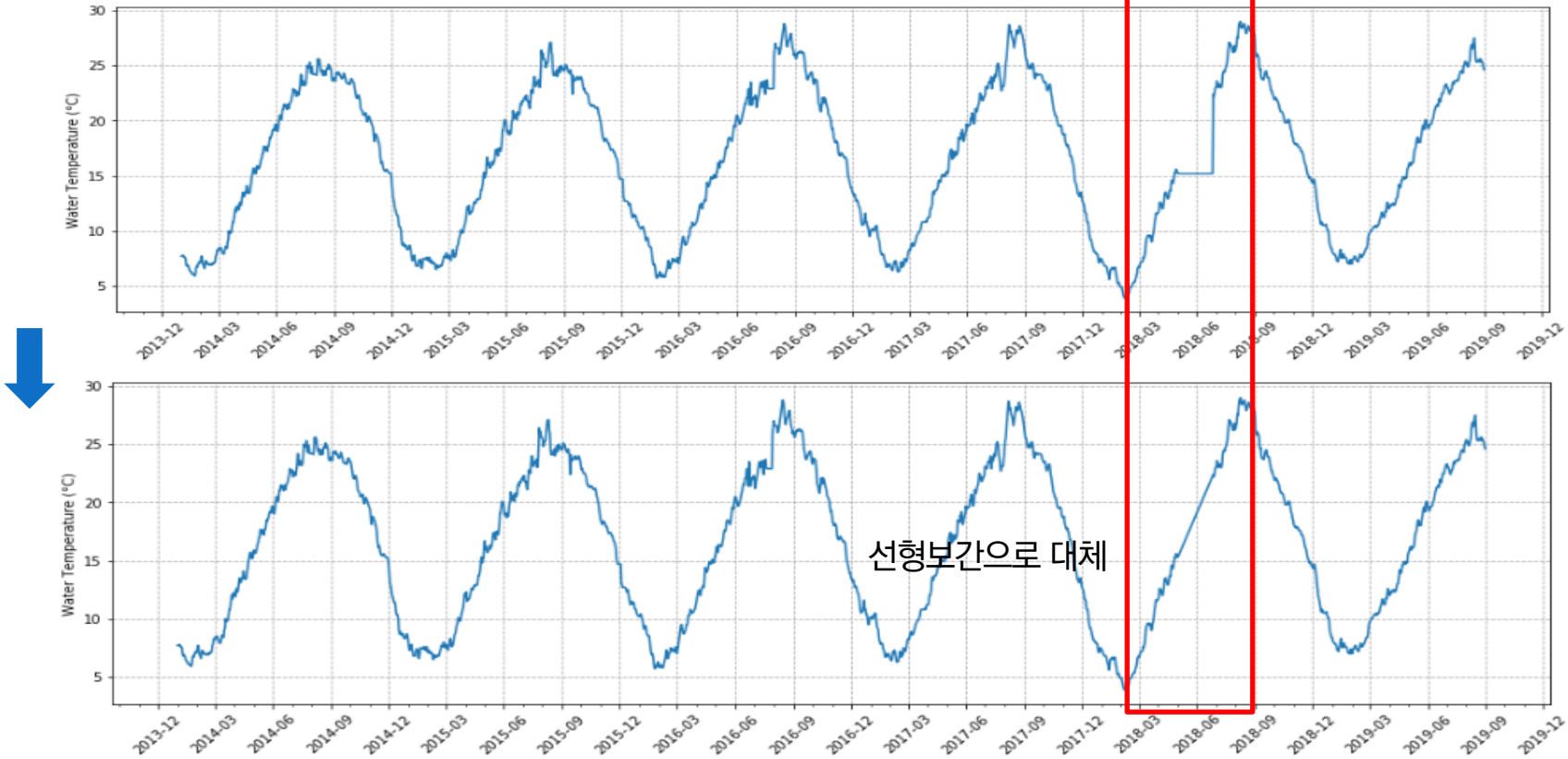
결측치 없음

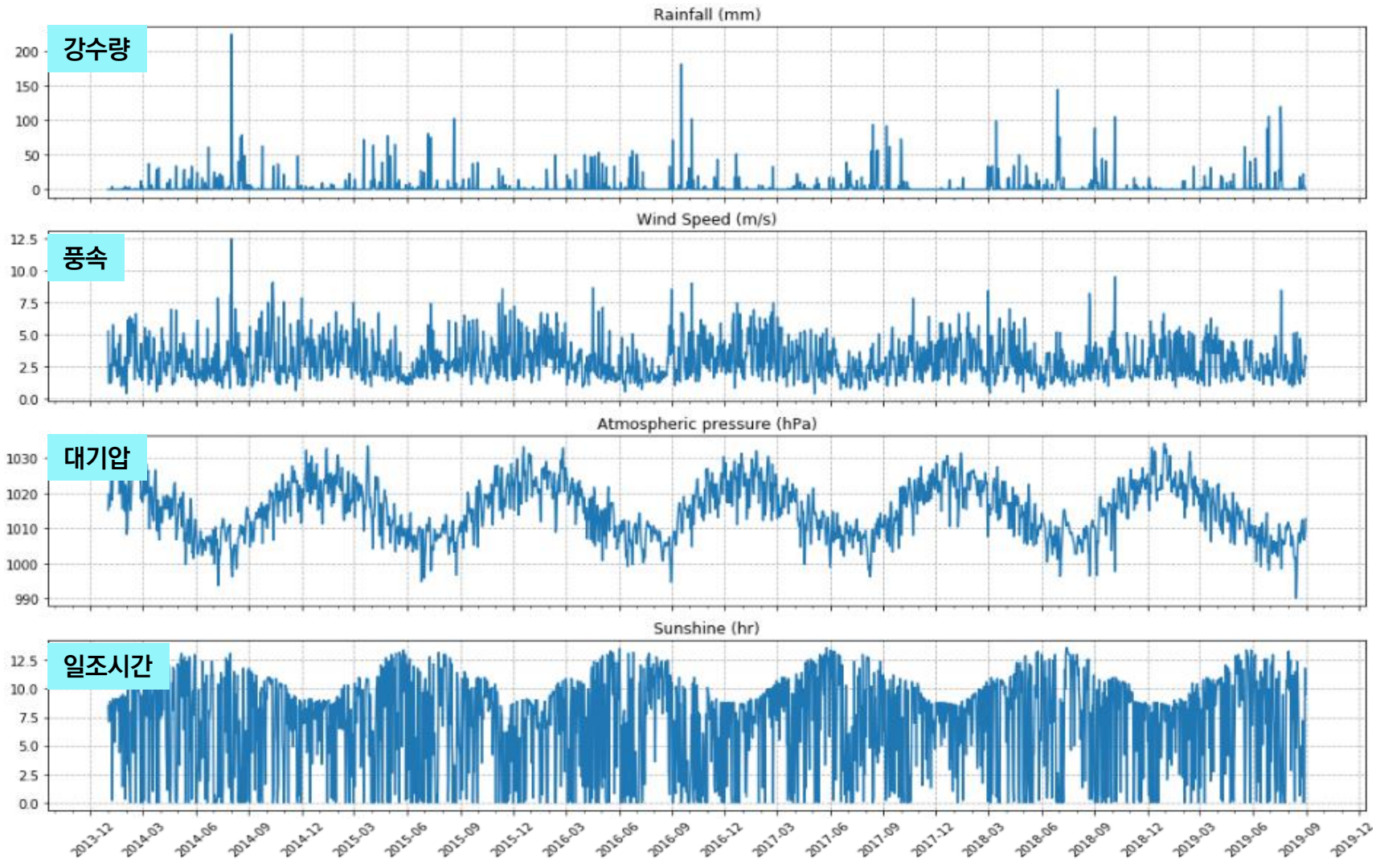
data info.

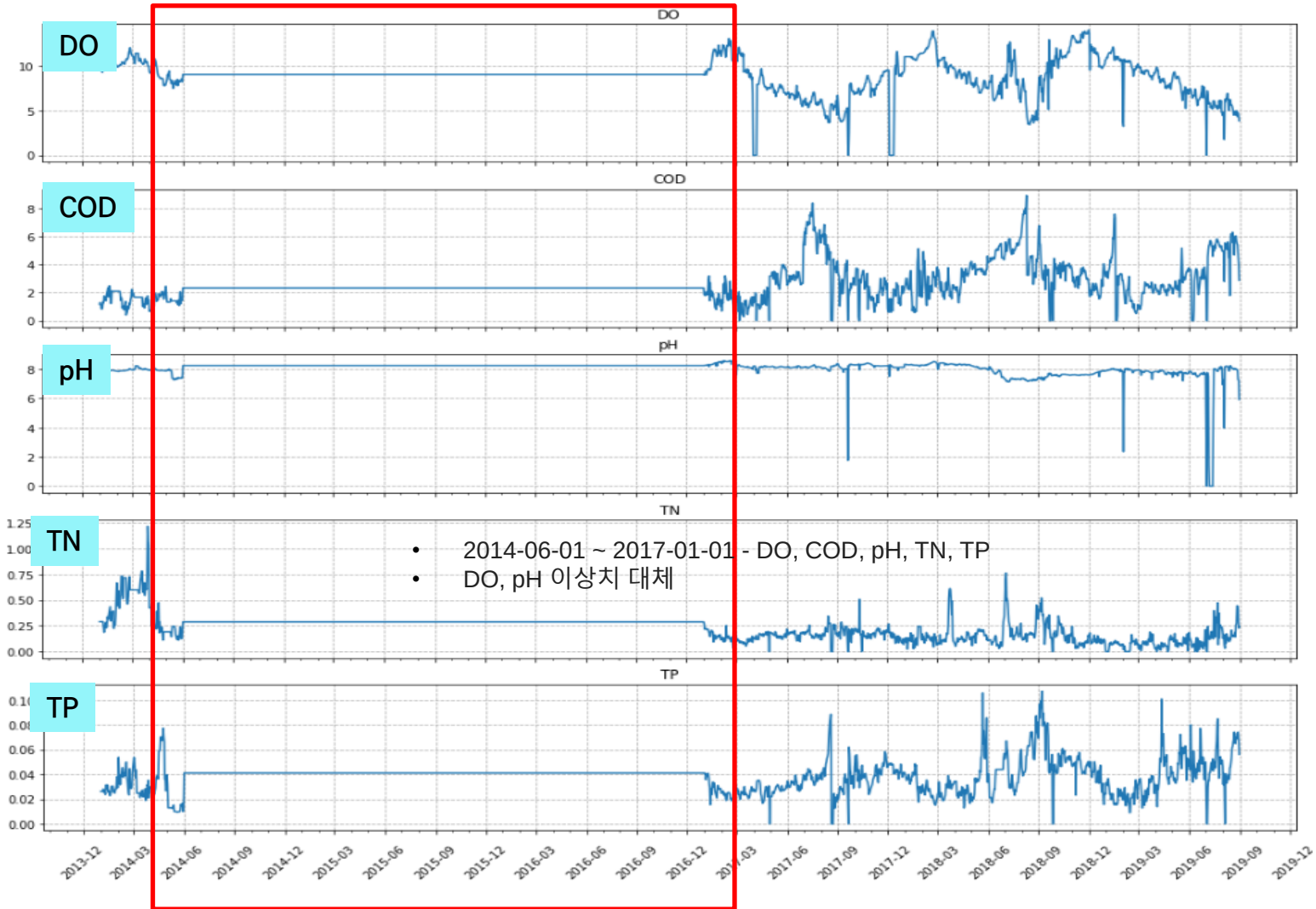
```
Data columns (total 14 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Date           2068 non-null   datetime64[ns]
 1   Temp           2068 non-null   float64
 2   Rainfall_Dur   2068 non-null   float64
 3   Rainfall_day   2068 non-null   float64
 4   WindSpeed      2068 non-null   float64
 5   Atm            2068 non-null   float64
 6   Sunshine       2068 non-null   float64
 7   DO             2068 non-null   float64
 8   COD            2068 non-null   float64
 9   pH             2068 non-null   float64
 10  TN             2068 non-null   float64
 11  TP             2068 non-null   float64
 12  Min_Density    2068 non-null   int64
 13  Max_Density    2068 non-null   int64
dtypes: datetime64[ns](1), float64(11), int64(2)
```

전처리 – 이상치, 오류데이터 처리

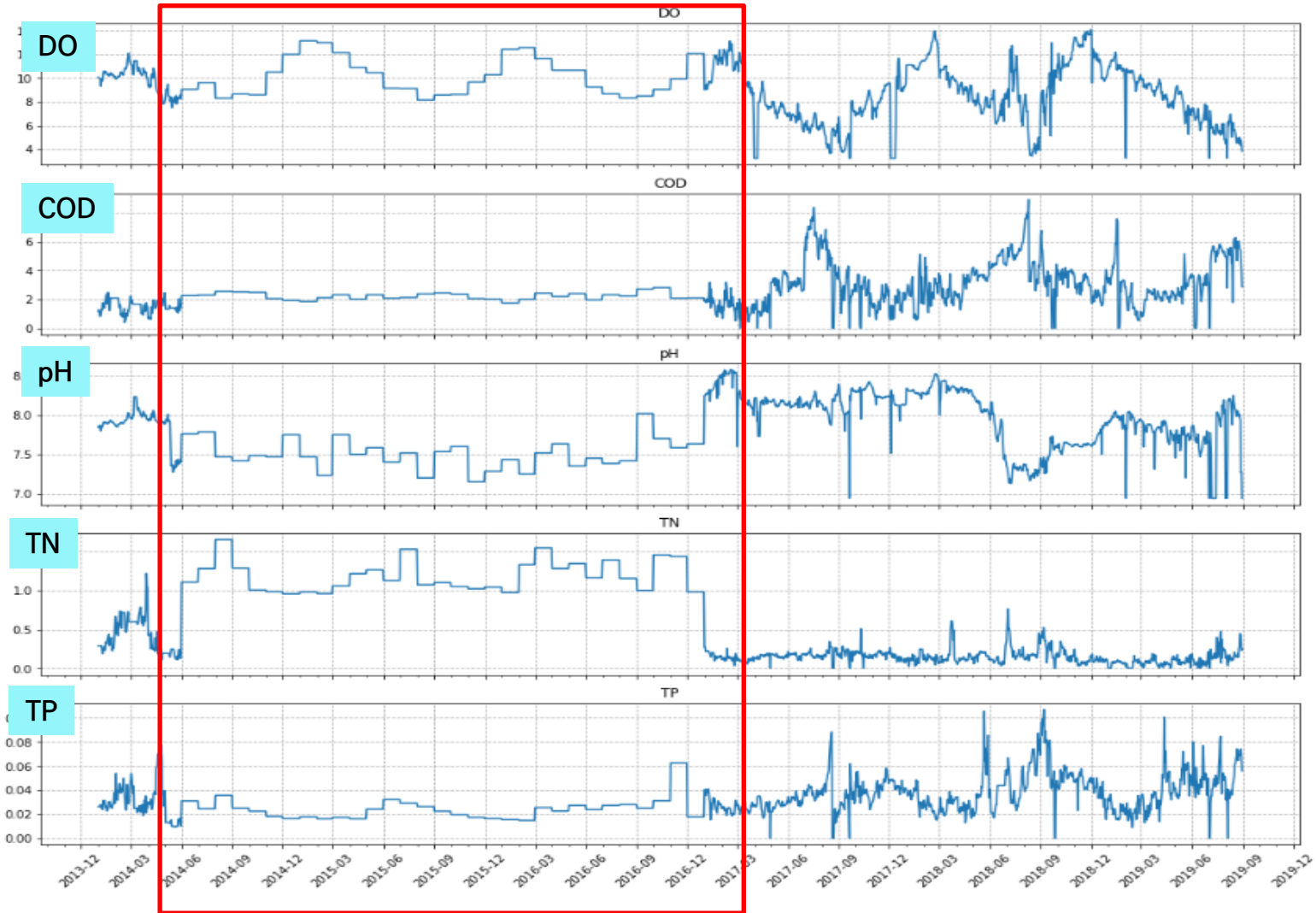
수온




Rainfall (mm)
강수량
200
150
100
50
0
Wind Speed (m/s)
풍속
12.5
10.0
7.5
5.0
2.5
0.0
Atmospheric pressure (hPa)
대기압
1030
1020
1010
1000
990
Sunshine (hr)
일조시간
12.5
10.0
7.5
5.0
2.5
0.0
2013-12
2014-03
2014-06
2014-09
2014-12
2015-03
2015-06
2015-09
2015-12
2016-03
2016-06
2016-09
2016-12
2017-03
2017-06
2017-09
2017-12
2018-03
2018-06
2018-09
2018-12
2019-03
2019-06
2019-09
2019-12



- 2014-06-01 ~ 2017-01-01 - DO, COD, pH, TN, TP
- DO, pH 이상치 대체


DO
COD
pH
TN
TP
DO
COD
pH
TN
TP
2013-12
2014-03
2014-06
2014-09
2014-12
2015-03
2015-06
2015-09
2015-12
2016-03
2016-06
2016-09
2016-12
2017-03
2017-06
2017-09
2017-12
2018-03
2018-06
2018-09
2018-12
2019-03
2019-06
2019-09
2019-12

가설 타당성에 대한 검토

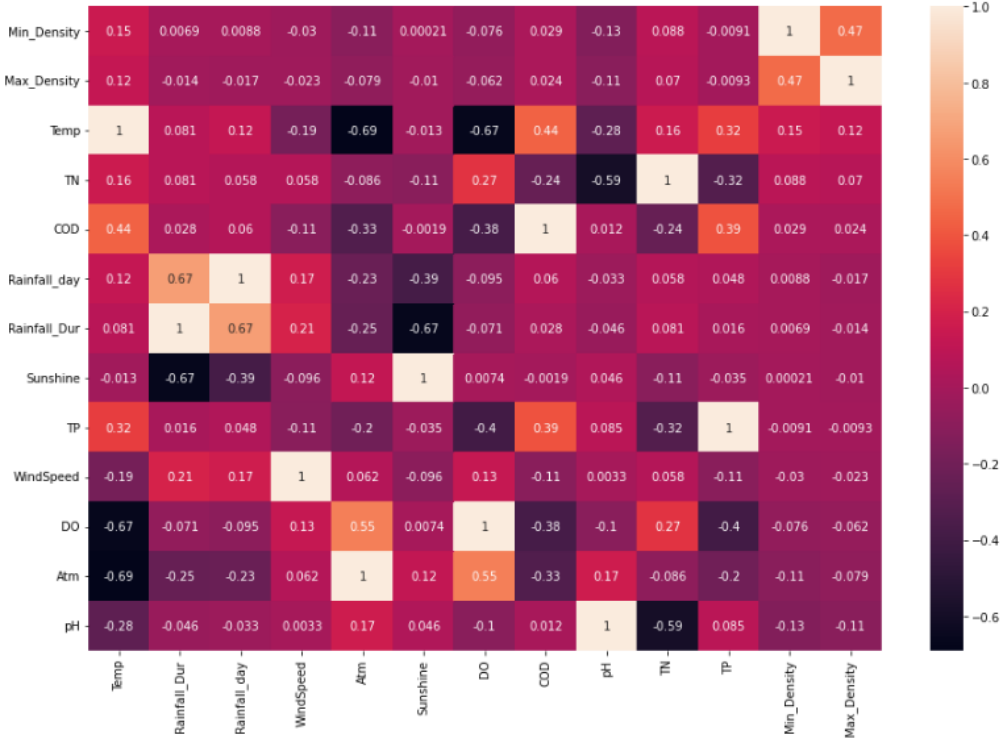
1. 여름철에 적조현상이 빈번하게 발생한다.

여름철 → 수온↑, 일조량↑, 강수량↑

2. 이 밖에 해양환경 등 적조 발생에 기여하는 요인이 있다.

▪ 특성간 양의 상관관계
강수지속시간 – 강수량 : 0.67
기압 – 용존산소 : 0.55
온도 – COD : 0.44

▪ 특성간 음의 상관관계
온도 – 기압 : −0.69
온도 – 용존산소 : −0.67
강우지속시간 – 일조시간 : −0.67
pH – TN : −0.59



- 히트맵상으로 확인해보았을 때 적조생물 밀도와 관계가 높은 특성은 Temp, TN, COD, Rainfall_day의 순으로 나타난다.
- 따라서 여름철과 적조현상은 높은 상관관계를 갖고 있을 것으로 보이며,
- 또한 계절요소 외에 TN, COD의 해양요소도 적조발생의 요인이 될 것으로 보인다.

특성공학

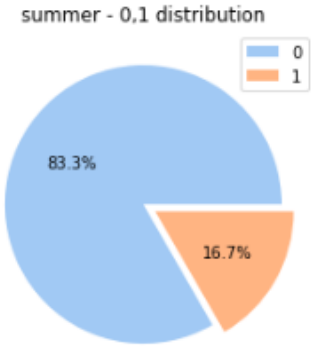
① 적조발생여부 특성 생성

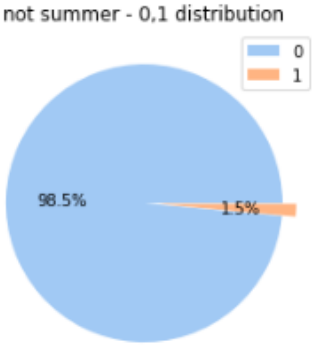
- 특성에 따른 적조발생 여부를 분석하기 위해, 적조생물 밀도 특성(max)을 *적조 주의보 발령기준*에 따라 0과 1로 구분한 적조발생여부 특성('Cochlo_yn') 생성
- 적조 주의보 발령기준 : Cochlodinium polykrikoides 100 이상

	Date	Temp	Rainfall_Dur	Rainfall_day	WindSpeed	Atm	Sunshine	DO	COD	pH	TN	TP	Cochlo_yn
0	2014-01-01	7.700	0.000	0.000	5.263	1,015.147	8.500	10.071	1.168	7.843	0.290	0.026	0
1	2014-01-02	7.700	0.000	0.000	1.219	1,019.767	7.100	10.001	1.297	7.859	0.290	0.026	0
2	2014-01-03	7.700	0.000	0.000	1.238	1,016.513	7.300	9.984	1.163	7.846	0.290	0.026	0
3	2014-01-04	7.800	0.000	0.000	2.419	1,016.210	8.900	9.946	0.845	7.850	0.290	0.026	0
4	2014-01-05	7.700	0.000	0.000	2.688	1,020.507	8.700	9.889	0.818	7.869	0.290	0.028	0

- 여름철(6~9월)과 비여름철로 데이터를 나누어 타겟값에 대한 분포 확인

→ 적조발생은 거의 여름철에 발생함을 확인할 수 있음





특성공학

② 직전 7일 평균값 특성 생성

- 적조발생 예측은 순간의 변수에 따라 타겟값이 결정되지 않으므로 직전 7일 평균값 특성 생성
- Date 기준으로 정렬후 Date 컬럼 drop

	Date	Cochlo_yn	m_Temp	m_Rainfall_Dur	m_Rainfall_day	m_WindSpeed	m_Atm	m_Sunshine	m_DO	m_COD	m_pH	m_TN	m_TP
0	2014-01-01	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	2014-01-02	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	2014-01-03	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	2014-01-04	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	2014-01-05	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	2014-01-06	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	2014-01-07	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	2014-01-08	0	7.686	0.000	0.000	2.242	1,019.814	7.971	9.878	1.075	7.846	0.290	0.026
8	2014-01-09	0	7.657	1.383	0.600	2.069	1,020.224	6.786	9.834	1.090	7.846	0.283	0.026
9	2014-01-10	0	7.557	1.383	0.600	2.718	1,020.926	7.071	9.808	1.155	7.846	0.268	0.026

	Cochlo_yn	m_Temp	m_Rainfall_Dur	m_Rainfall_day	m_WindSpeed	m_Atm	m_Sunshine	m_DO	m_COD	m_pH	m_TN	m_TP
0	0	7.685714	0.000000	0.0	2.241964	1019.813571	7.971429	9.877769	1.075438	7.845976	0.290400	0.026396
1	0	7.657143	1.382857	0.6	2.068750	1020.223631	6.785714	9.834114	1.090010	7.846190	0.283286	0.026392
2	0	7.557143	1.382857	0.6	2.717798	1020.925714	7.071429	9.808460	1.154867	7.846048	0.268414	0.026043

04

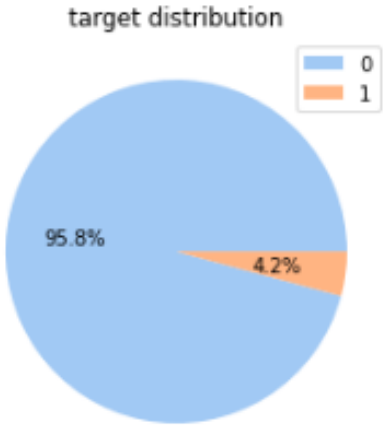
데이터 분석 및 예측

- 모델링
- 모델해석

모델링

① 타겟분포 확인

- 불균형 클래스이므로 모델적용시 **class_weight** 조절



② 데이터셋 분할

- train, val, test셋으로 분할 & 타겟('Cochlo_yn'), 특성 분리
  - 모델학습시 train, 모델 성능 검증시 validation, 최종모델 성능 검증시 test셋 사용

③ 기준모델 설정

- 로지스틱회귀모델로 하이퍼파라미터를 조절하지 않은 모델

```
훈련 Accuracy Score: 0.9602
검증 Accuracy Score: 0.9576
검증 F1 Score: 0.2222
검증 AUC Score: 0.9476
              precision    recall  f1-score   support

           0       0.96      0.99      0.98       158
           1       0.50      0.14      0.22         7

    accuracy                           0.96       165
   macro avg       0.73      0.57      0.60       165
weighted avg       0.94      0.96      0.95       165
```

◆ 분류모델 성능지표

- 정확도 $(Accuracy) = \frac{TP + TN}{TP + FN + FP + TN}$
- 정밀도 $(Precision) = \frac{TP}{TP + FP}$
- 재현율 $(Recall) = \frac{TP}{TP + FN}$
- AUC : 임계값에 따른 TRP, FPR 그래프의 아래 면적

모델링

④ 로지스틱 회귀모델

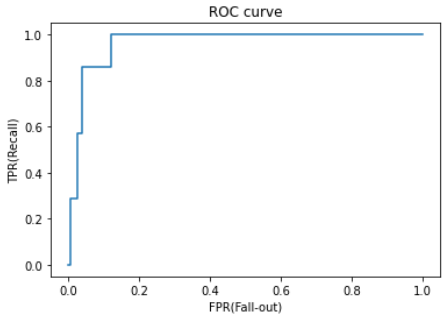
- 최적 하이퍼파라미터: {'solver': 'liblinear', 'penalty': 'l1', 'max_iter': 100, 'C': 1}

```
Precision Score:  0.2857
Recall Score:  0.8571
F1 Score: 0.4286
Accuracy Score: 0.903
AUC Score: 0.9629
```

→

```
Precision Score:  0.2593
Recall Score:  1.0
F1 Score: 0.4118
Accuracy Score: 0.8788
AUC Score: 0.9629
```

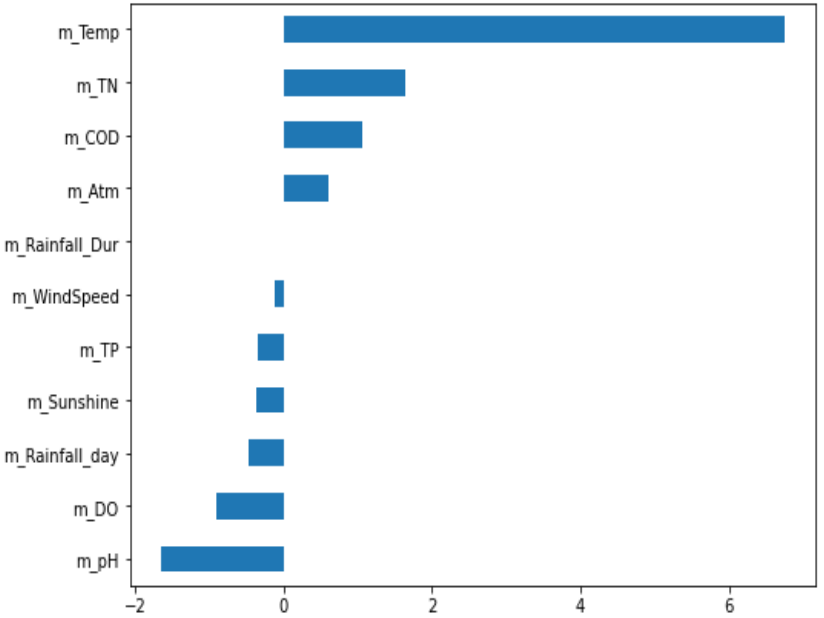
최적 임계값 0.4로 조정



- 다른 지표는 떨어졌지만 recall 값은 0.1429 상승

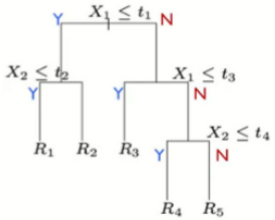
※ 최적 임계값 :
TRP – FPR 이 최대가 되는 지점

회귀계수 해석



- 회귀계수를 통해 특성들이 타겟값에 미치는 영향을 확인
- 수온, TN, COD, 기압 순으로 해당 값이 커질수록 적조가 발생할 수 있다고 해석할 수 있음

모델링



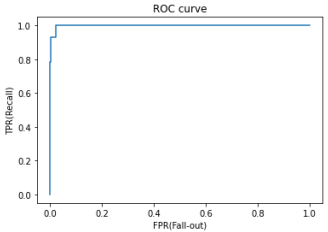
⑤ 랜덤포레스트

– 최적 하이퍼파라미터: {'max_depth': 15, 'max_samples': 1.0,
'min_samples_leaf': 1, 'n_estimators': 206}

```
Precision Score:  0.9167
Recall Score:  0.7857
F1 Score: 0.8462
Accuracy Score: 0.9879
AUC Score: 0.998
```

```
Precision Score:  0.9286
Recall Score:  0.9286
F1 Score: 0.9286
Accuracy Score: 0.9939
AUC Score: 0.998
```

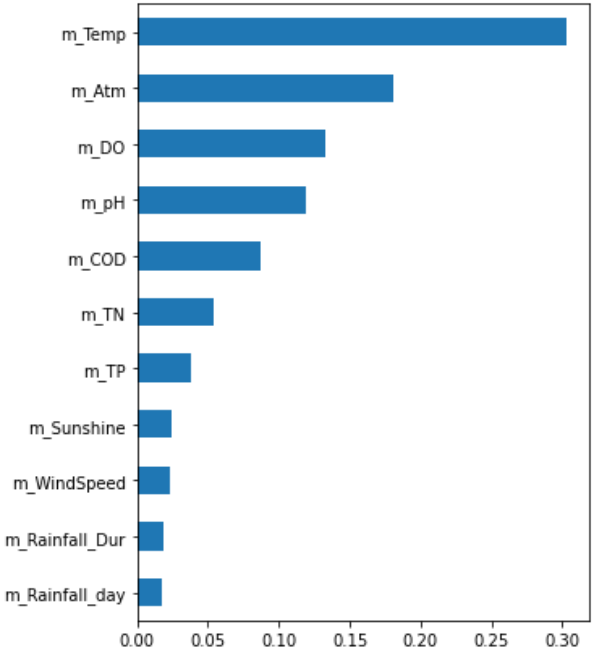
최적 임계값 0.4로 조정



랜덤포레스트 모델의 성능이 더 높은 이유?

- 로지스틱회귀분석은 각 독립적인 변수들이(독립변수) 종속변수와 선형관계에 있다고 가정
- 랜덤포레스트는 변수간 상호작용이 있거나 비선형적 관계의 데이터에도 잘 적용됨

Feature Importance(MDI) 확인



- 카테고리형이 있는 데이터셋이 아니고 특성간 상호작용이 존재하므로 mdi(평균불순도감소) 특성중요도를 사용
- 수온, 기압, 용존산소, pH 등의 순으로 해당 특성이 예측값에 영향을 미침

모델링

⑥ 최종모델 일반화 성능

```
Precision Score:  0.85
Recall Score:  1.0
F1 Score: 0.9189
Accuracy Score: 0.9927
AUC Score: 0.9979
```

기준모델 성능

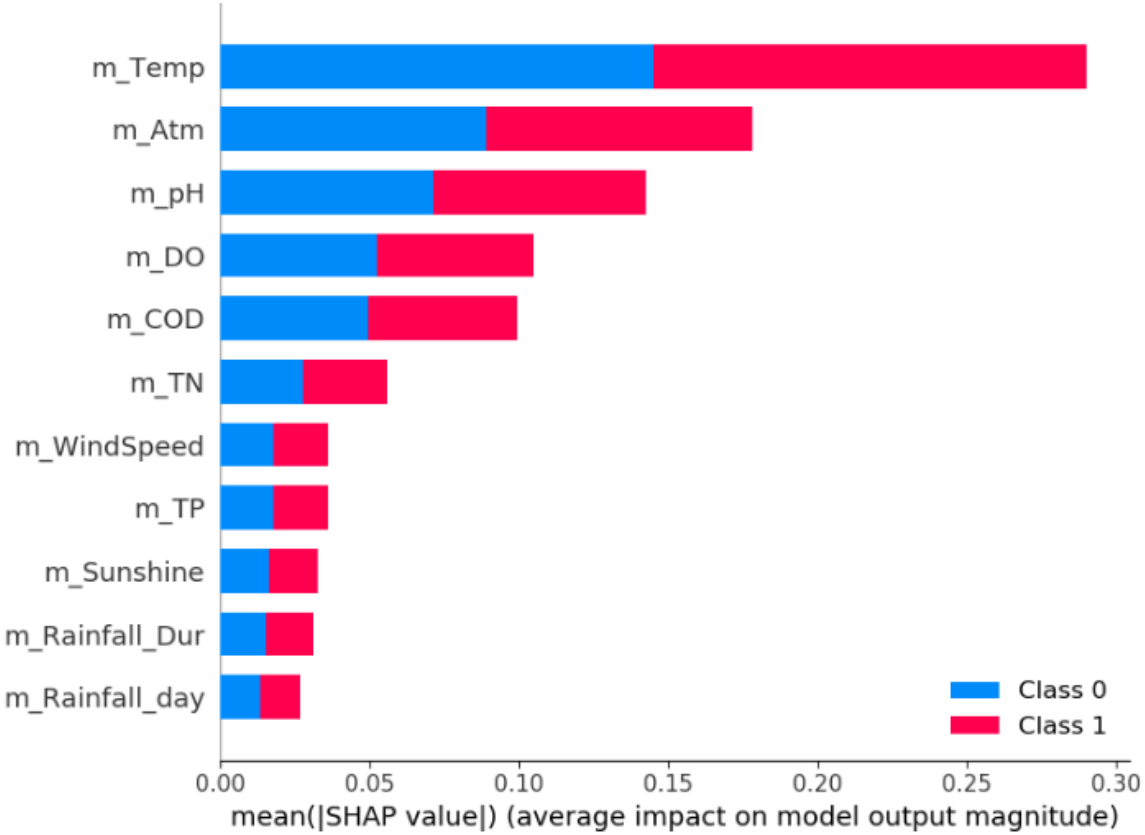
```
훈련 Accuracy Score: 0.9602
검증 Accuracy Score: 0.9576
검증 F1 Score: 0.2222
검증 AUC Score: 0.9476
              precision    recall  f1-score   support

           0       0.96      0.99      0.98       158
           1       0.50      0.14      0.22         7

    accuracy                           0.96       165
   macro avg       0.73      0.57      0.60       165
weighted avg       0.94      0.96      0.95       165
```

모델 해석

❖ SHAP


m_Temp
m_Atm
m_pH
m_DO
m_COD
m_TN
m_WindSpeed
m_TP
m_Sunshine
m_Rainfall_Dur
m_Rainfall_day
Class 0
Class 1
0.00
0.05
0.10
0.15
0.20
0.25
0.30
mean(|SHAP value|) (average impact on model output magnitude)

Thanks