● 2023 DATA·AI 분석 경진대회 사과 품질 예측 모델 개발

아삭 팀

Contents

 01

 프로젝트 개요

02

활용 데이터

03

모델 개발 방법

04

실험 및 평가

05

프로그램 개발

06

활용 계획 및 기대효과

07

시연

01 프로젝트 개요



기후 변화의 영향을 가장 많이 받는 과일은 사과로, 오는 2050년대에는 강원도 일부 산지에서만 사과를 재배할 수 있을 것으로 예측한다.



by 국립원예특작과학원 온난화대응농업연구소



사과는 우리나라 과일 재배 농가의 16.8%를 차지하고, 재배 면적도 가장 넓은 작목이다. 그러나 기후변화로 인해 사과의 재배 지역이 북상하고 있으며, 2100년에는 사과가 강원도 일부에서만 재배될 것이라는 관측도 나온다.



by 농촌진흥청 국립원예특작과학원장



기후 변화에 맞춰 효율적으로 사과를 생산하기 위한 대응 방안의 필요성을 느껴 해당 주제로 연구를 진행

┃ 01 프로젝트 개요



★ 사과 생산에 영향을 미치는 기상 요인 분석

사과의 품질(크기, 단맛, 쓴맛 등)은 어느 시기(특정 계절, 월, 일 등)의 어떤 기상요소(온도, 일사, 강수량 등)가 가장 큰 영향을 미치는 지 분석



★ 사과 품질 예측 모델 개발

분석한 기상 요인을 바탕으로 머신러닝을 활용하여 사과의 품질을 예측해주는 모델 개발

[품종에 따른 개별 모델 최적화]

사과는 품종에 따라 성장 시기 및 특징이 다르다. 따라서, 우리나라의 대표적인 사과 품종인 홍로와 후지를 선정하고, 품종별 특징을 고려하여 품질을 예측할 수 있도록 개별 모델을 생성한다.

• 홍로: 8월 하 ~ 9월 상, 중에 재배, 산이 덜한 것이 특징

후지: 10월 하~ 11월 상에 재배, 당과 산의 적절한 조화가 특징

[사용자 맞춤형 사과 품질 예측]

사용자가 직접 1년 동안의 기상 데이터를 준비하기는 어렵다. 따라서, 지점과 품종, 수확일을 입력할 경우 API를 통해 자동으로 기상 데이터를 불러오며 품질을 예측해주는 모델을 생성한다.







1. target data

무게, 당도에 대한 농촌진흥청의 사과 품질 기준과 당산비에 대한 기사를 참조하여 품질 기준 선정

- 사과 품질 데이터 중에서 과중, 당도, 당산비를 이용하여 품질 등급 생성
- 이때, 품종을 고려하여 품질 등급 기준을 생성하며 1등급부터 4등급까지 총 4등급으로 구성

과중 / 당도

by 농촌진흥청

구분	농산물 품	농촌진흥청			
항목	E	상	'Top fruit'		
무게	후지 : 250g이상 홍로 : 250g이상	후지 : 215g이상 홍로 : 215g이상	홍로 : 320g±10%		
색택	품종고유색택 후지 : 60%이상 홍로 : 70%이상	품종고유색택 후지 : 40%이상 홍로 : 50%이상	품종고유색택 후지 : 70%이상 홍로 : 80%이상		
당도	후지 : 14°Bx이상 홍로 : 14°Bx이상	후지 : 12°Bx이상 홍로 : 12°Bx이상	후지 : 14°Bx이상 홍로 : 14°Bx이상		

• 과중 : 320g 이상 / 250 이상 / 215 이상 / 215 미만

• 당도: 14°BX 이상 / 12°BX 이상 / 12°BX 이상 / 12°BX 미만

당산비



과일 맛 결정하는 '황금비'

떫은맛은 사라지고 단맛과 신맛은 절묘한 조화

by 동아사이언스

하지만 달기만 하다고 무조건 맛있는 과일은 아니다. 설탕을 한 스푼 가득 떠먹었을 때 맛있다고 느끼지 않는 것과 마찬가지다. 단맛과 신맛이 적당히 조화를 이뤄야 좋은 맛을 낸다. 이런 황금비율을 당산비(당함량/산함량)라 한다. 사과의 경우 가장 달다고 느끼는 종은 당산비가 41 이상이다. 우리나라에서 가장 많이 재배되는 사과인 후지는 당산비가 36 정도다.

• 당산비: 41 이상 / 36 이상 / 36 미만



위 3가지 품질 기준을 동시에 만족하는 경우, 해당 등급을 부여

1. target data

무게, 당도에 대한 농촌진흥청의 사과 품질 기준과 당산비에 대한 기사를 참조하여 품질 기준 선정

- 사과 품질 데이터 중에서 과중, 당도, 당산비를 이용하여 품질 등급 생성
- 이때, 품종을 고려하여 품질 등급 기준을 생성하며 1등급부터 4등급까지 총 4등급으로 구성

품질 등급 기준

	1등급	2등급	3등급	4등급
과중	320g 이상	250g 이상	215g 이상	215 미만
당도	14°BX 이상	12°BX 이상	12°BX 이상	12°BX 미만
당산비	41 이상	36 이상	36 🗆	미만

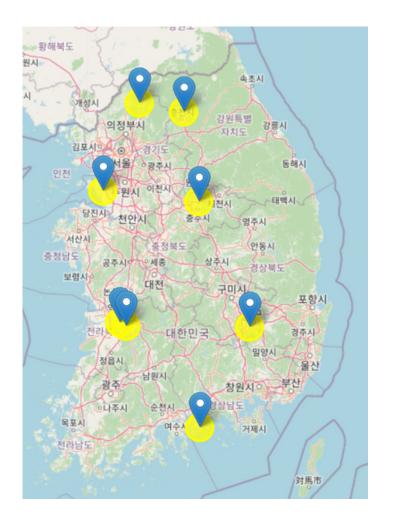


2. feature data

제공 받은 사과 품질 데이터의 지점과 날짜 이용하여, 기상청에서 2016년부터의 기상데이터 수집

- 기상 지점은 전국 8개 대표 과수원과 가장 가까운 위치의 지역을 선택
- 기상 데이터는 사과 생산에 영향을 미칠 것이라 판단되는 10가지 데이터로 구성했으며, 월 단위로 가져옴

과원 위치와 기상 지점



<u> 과원 위치</u>		<u>기상 지점</u>
- 화 보 시 사 사 수 수 수 수	$\begin{array}{ccc} \rightarrow & \\ \end{array}$	수원 동두천 전주 남해 충주 북춘척
군위	\rightarrow	대구

기상 데이터



2. feature data

기상청에서 불러온 데이터프레임은 지점과 일시에 해당하는 기상 데이터 BUT, 우리가 필요한 데이터는 사과가 자라온 환경에 대한 기상 데이터

➡ 조사일자로부터 12개월 전의 기상 데이터를 칼럼으로 가지는 데이터프레임으로 변경

지점명	일시	평균기온(°C)	평균최고기온(°C)	평균최저기온(°C)	최고기온(°C)	 합계 일사량(MJ/m2)	평균지면온도(°C)
동두천	Jan-22	-1.1	13.3	1	22.2	 258.11	6.6
동두천	Feb-22	-3.9	5	-6.1	12.8	 248.82	-0.5



조사일	과원이름	1월 평균 기온(°C)	2월 평균 기온(°C)	 12월 평균 기온(°C)	 1월 평균 지면온도 (°C)	 12월 평균 지면온도(°C)
2022-10-27	포천	-3.9	-2.6	 -1.1	 -3.2	 -0.5

※ 조사일자가 15일 이전이라면 => 해당 월의 기상데이터 미포함

※ 조사일자가 15일 이후라면 => 해당 월의 기상데이터 포함

데이터프레임 예시

조사일	과원이름	1월 평균 기온(°C)	2월 평균 기온(°C)	 12월 평균 기온(°C)	 1월 평균 지면온도 (°C)	 12월 평균 지면온도(°C)	품질
2022-10-27	포천	-3.9	-2.6	 -1.1	 -3.2	 -0.5	2등급
2022-10-26	화성	-2.2	-1.1	 0.7	 -1.7	 0.7	3등급

➡ 총 122개의 독립 변수와 1개의 종속 변수를 가진다.

칼럼의 수가 너무 많아 사과 생산에 영향을 미치는 변수를 분석하는데 어려움이 존재 따라서, 모델링에 앞서 **상관분석**과 **요인분석**을 통해 중요한 칼럼을 선별하는 과정을 수행 ※ 후지와 홍로 모두 각각 상관분석/요인분석을 진행

데이터 수집

데이터 분석

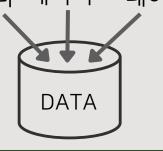
데이터 전처리

모델링

프로그램 생성

데이터 수집

데이터 데이터 데이터



사과 품질 결정

과중

당도

당산비

요인분석, 상관분석

품종에 따라 요인분석과 상관분석을 각각 진행

요인분석의 경우 새로운 변수를 생성

새로운 변수 추가

- 태풍 횟수
- 태풍평균강도
- 장마일수
- 합계강수량
- 폭염일수

모델링

- 랜덤포레스트
- xgboost
- adaboost
- knn
- gradientboosting

앙상블

예측 평가를 통해

성능 기반 모델 선택

요인 분석 선택

SHAP 분석



기상청 추가 기상



EDA 분석



데이터 전처리

라벨 인코딩

결측치 제거

변수 Scaling

프로그램



사용자 지점, 품종,

수확일 입력



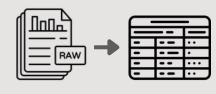
사과 품질 예측 모델



[예측 결과]

- 사과 품질
- 품질 확률
- 지도 시각화
- 중요 요인 시각화

데이터 프레임 변환



원하는 형태로 변경

1. 상관분석

사과 품질 등급과 각 월별 기상요인의 상관분석을 기초로, 상관계수를 이용하여 변수 선정

- 상관계수는 두 연속변수의 선형관계 정도를 나타내는 척도로, -1과 1의 사이 값으로 나타낸다.
- 품질 변수는 연속형 데이터가 아니므로, 순위형 변수의 상관계수를 구하는 스피어만 상관계수를 사용한다.

후지 품종

11월 일조율(%) -	0.37
5월 월합강수량(00~24h만)(mm)	0.34
9월 월합강수량(00~24h만)(mm) -	0.3
9월 평균 상대습도(%) -	0.3
6월 평균 상대습도(%) -	0.27
3월 월합강수량(00~24h만)(mm)	0.27
3월 평균 최고기온(°C) -	0.26
10월 일조율(%) -	0.26
12월 합계 일사량(MJ/m2) -	0.26
12월 일조율(%) -	0.24
3월 평균 기온(°C) -	0.23
5월 평균 상대습도(%) -	0.22

홍로 품종

_		
	9월 월합강수량(00~24h만)(mm) -	0.25
	1월 일조율(%) -	0.18
	12월 일조율(%) -	0.18
	10월 월합강수량(00~24h만)(mm)	0.17
	7월 월합강수량(00~24h만)(mm) -	0.14
	1월 평균 최고기온(°C) ·	0.11
	11월 평균 최고기온(°C) -	0.1
	11월 월합강수량(00~24h만)(mm)	0.096
	5월 일조율(%) -	0.088
	1월 합계 일사량(MJ/m2) ·	0.078
	2월 월합강수량(00~24h만)(mm) -	0.078
	4월 합계 일사량(MJ/m2) -	0.077
		- 품질

2. 요인분석

상관관계를 이용하여 변수들 간의 체계적인 구조를 밝히고 서로 유사한 변수들끼리 묶어주는 통계 기법 즉, 여러 개의 변수들이 어떻게 연결되어 있는가를 분석하여 변수들 간의 관계를 <mark>공통 요인</mark>으로 설명



※ 요인분석은 품질 변수와 관계 없이 서로 유사한 독립 변수들을 묶어주는 기법이므로, 후지와 홍로의 요인분석 결과는 동일하다.

2. 요인분석

1월 평균 기온(*C)	0.95	0.06	-0.33	-0.12	0.03	0.15	-0.02	-0.09	0.01	0.09	-0.00	-0.07	0.02	-0.14
2월 평균 기온(°C)	0.96	-0,01	-0.09	0,23	-0.03	0.12	-0,06	-0,15	-0,00	0.00	-0,34	-0.17	0.23	-0.14
3월 평균 기온(°C)	0,62	-0,01	0,21	0,19	-0.05	0,22	0,10	0,10	0,02	-0,27	0,13	-0,25	0.09	-0,01
4월 평균 기온(°C)	0.49	0.05	0.78	-0.05	-0.16	-0.05	-0.02	-0.07	0.01	0.12	-0.14	-0.02	0.02	0.10
5월 평균 기온(°C)	0.45	0.02	0.01	-0,56	-0.27	0.07	0.22	0,34	0.02	0.06	0,02	0.02	0.04	-0.03
6월 평균 기온(°C)	0,21	-0,02	0,12	-0,11	-0,30	-0,20	0,87	0,21	0,02	-0,24	0,13	-0,03	0,17	0,01
7월 평균 기온(°C)	0.05	0.02	1,08		0.06	0.05		0,02	0,01	0.24	-0,03	-0,04	0.26	0,03
8월 평균 기온(°C)	0.17	0.00	0.03	-0.00	-0.03	0.04	0,37	0.79	0.01	-0.02	0,35	-0.08	0.18	-0.01
9월 평균 기온(°C)	0.64	-0.09	0.29	-0,07	80.0	-0.00	0.05	-0,01	0.02	-0.10	0.05	-0.08	0.66	-0.02
10월 평균 기온(°C)	0,89	0,05	0,19	-0,00	-0,01	0.11	0,00	-0,06	-0,00	0.27	0,02	0,17	0.07	0.01
11월 평균 기온(°C)	0.91	0.06	-0.02	-0.01	-0.09	-0.18	-0.18	-0.01	0.00	-0.09	-0.04	-0.03	0.07	-0.03
12월 평균 기온(°C)	0.81	0.07	0.02	-0,13	0.13	-0,21	-0,00	0.04	0.01	0.24		-0.19	0.04	0.16
1월 평균 최고기온(°C)	0,84	-0,08	-0,31	-0,26	0,16	0,26	0,04	-0,11	-0,02	0,04	-0,08	-0,17	0.05	-0,05
2월 평균 최고기온(°C)	0.85	-0.12	-0.03	0,23	0.06	0.18	0.06	-0.11	-0.03	-0.02	-0.43	-0.27	0.18	-0.07
3월 평균 최고기온(°C)	0.46	-0.12	0.24	0.13	-0.02	0.22	0.29	0.14	-0.03	-0.33	-0.09	-0.32	0.01	-0.14
					:		:			:				
										•				
4월 평균지면온도(°C)														0,26
5월 평균지면온도(°C)	0,26	-0.03	-0.11	-0.84	-0.01	0,13	0,14	0.05	0.00	0.02	0,19	0,18	-0.18	0,18
6월 평균지면온도(°C)	-0.03	-0.04	0.02	-0.23	-0.10	0.11	0.63	0.28	0.00	0.19	-0.08	0.25	-0.10	0.13
7월 평균지면온도(°C)	-0.14	0.00	0.90	0,19	0,13	0,14	0.06	0.25	-0,01	0.14	0.04	0.04	-0.06	0.07
8월 평균지면온도(°C)	-0.06	-0.02		0.07	0.04	-0.01	0.14	0.95	0.05	0.03	0.16	0.04	-0.15	0.03
9월 평균지면온도(°C)	0.54	-0.08	0.22	-0.22	0,10	-0.02	0.07	0.03	0.01	-0.15	-0.07	0,25	0,32	0.06
10월 평균지면온도(°C)	0,88	-0,02	0.15	-0,03	0,01	0,11	0.04	-0.10	0.01		-0,02	0,22	-0.08	0.07
11월 평균지면온도(°C)	0.87	0.01	-0.01	0.02	-0.04	-0.24	-0.15	0.06	-0.02	-0.11	-0.09	0.03	-0.06	-0.01
12월 평균지면온도(°C)	0.78	0.04	0.01	-0,13	0,17	-0.36	-0.03	0.02	-0,01	0.23	0.07	-0.14	-0.01	0.09
	Ĺ.	17	5	47	72	91	17	0	ō	10	=	12	13	4
	Factor'	Factorí	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9	actor10	actor1	actor	actor	actor14



각 변수와 요인 간의 상관관계 정도를 나타내며, 값이 높은 변수가 해당 요인에서 중요한 변수

<u>요인 해석</u>

- 각 factor별로 변수들의 공통된 특성 파악
- 공통된 특성을 바탕으로 요인에 대해 요인명 부여



<u>경우 1</u>

특정 월에 대한 기상 요인들의 모임

ex) factor 8 => 8월 온도

8월 평균 최고기온, 8월 일조율, 8월 평균 기온, 8월 평균지면온도



<u>경우 2</u>

특정 기상 요인에 대한 월들의 모임

ex) factor2 => 평균풍속

1, 2, 3, 4, 5, 11, 12월의 평균풍속

2. 요인분석

요인 해석 결과

factor1	1_2월의 온도	factor8	8월의 온도
factor2	평균풍속	factor9	합계일사량
factor3	7월의 기온	factor10	10월_12월 월합강수량
factor4	5월의 온도	factor11	기타
factor5	평균 상대 습도	factor12	9월의 일조율
factor6	11월_12월 일조율	factor13	9월의 온도
factor7	6월의 기온	factor14	12_1_2_4월 일조율

3. 추가 변수 - 태풍

'태풍 피해일자와 최대풍속에 따른 사과 낙과 피해율 분석' 논문에 따르면

최근 지구온난화와 같은 기후변화로 인해 강력한 태풍의 발생 비율이 증가하고 있으며, 특히 7월에서 10월 사이에 주로 발생하고 있어 사과 수확량 감소 및 품질 저하를 야기한다.

태풍 선정 기준

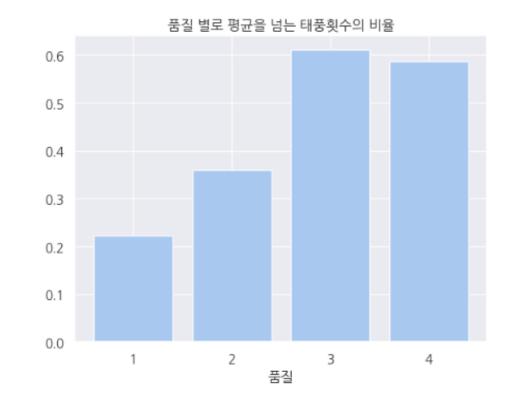
과실의 낙과 피해는 무게에 비례하기 때문에, 과실의 무게가 큰 수확기 때의 피해가 크다. 따라서, 해당년도의 태풍만 고려

=> 지금까지 기상데이터는 월별 데이터였지만 태풍 칼럼은 연도별 데이터로 입력된다.

우리나라에 영향을 미쳤다고 판단된 태풍으로, 영향도가 직접 영향이거나 상륙인 태풍만 고려

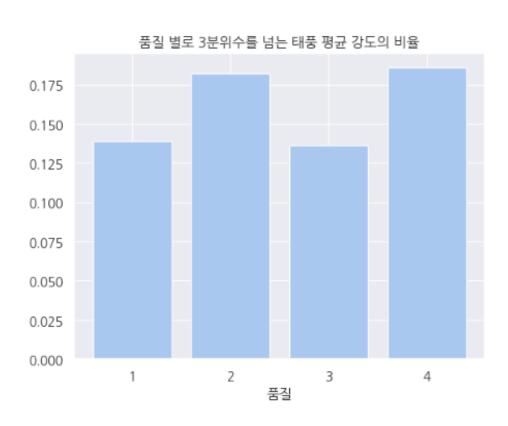
태풍횟수 변수

• 연도별 태풍 발생 횟수를 나타낸 변수



태풍평균강도 변수

• 태풍 평균 강도 = 각 태풍의 최대 풍속의 합 /발생 횟수



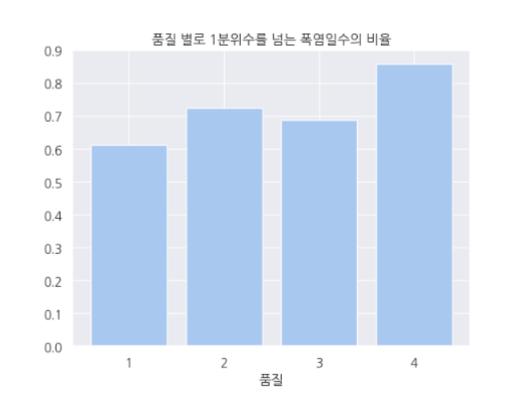
3. 추가 변수 - 폭염과 장마

폭염과 긴 장마는 습하고 더운 환경을 유발하기 때문에 균에 약한 사과에 영향을 주어 탄저병을 유발한다. 또한, 장마가 지속된다면 일조량이 부족해져 사과가 제대로 익지 못하는 문제가 발생하게 된다.

→ 결과적으로, 과도한 폭염과 장마는 사과 수확량 감소 및 품질 저하를 야기한다.

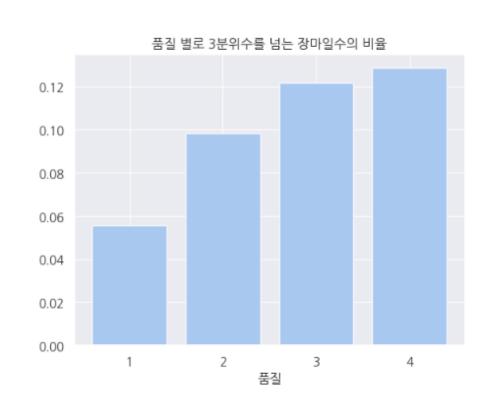
폭염 일수

• 지역 별로 각 수확년도의 체감온도가 33도 이상이었던 날의 횟수를 나타낸 변수



장마 일수 / 합계강수량

- 장마일수: 지역별로 장마가 지속된 기간을 나타낸 변수
- 합계강수량: 지역별로 해당 장마 기간 동안 내린 합계 강수량을 나타낸 변수



장마 일수는 짧음에도 불구하고, 합계 강수량이 많은 경우를 고려해야하므로

=> 장마와 관련하여 합계강수량 변수도 추가

4. 데이터 전처리

● 라벨인코딩

- target 변수가 범주형 데이터이므로 수치형 데이터로 변환
- '품질' 칼럼에 대해서만 진행
- '1등급' -> 0, '2등급' -> 1, '3등급' -> 2, '4등급' -> 3

● 결측치 처리

- 결측치나 이상치가 있는지 확인
- 기상 데이터의 결측치는 평균으로 대체
- fillna 메서드 사용

● 스케일링

- 데이터의 단위나 척도가 서로 다르므로 스케일링 필요
- 모든 X 변수들에 대해 진행
- StandardScaler 사용해 각 특성별 표준화

[후지 - 상관분석]

		<u>X train</u>											
		-010141-11-1						-11 == +1 .		-1-1011	+1-11-1 1 -1		
115	월 일조율	5월 월합강수량	9월 월합강수량	9월 평균 상대습도	•••	12월 합계 일사량	12월 일조율	태풍횟수	•••	장마일수	합계강수량	품질	
-0	.1686	-1.1068	0.7352	0.6279		0.4522	0.4812	0.4663		0.5007	-0.5102	2	
0	.6894	0.2416	-0.3433	1.0336		0.8437	-0.5487	0.4663		-1.0878	-0.4077	1	

<u> 상관분석을 통해 선별한 10개의 기상 변수</u>

<u>요인분석을 통해 생성한 14개의 기상 변수</u>

<u>추가 변수</u>

<u>추가 변수</u>

[후지 - 요인분석]

	<u>X train</u>													
1_	2월 온도	평균풍속	7월의 기온	5월의 기온		9월 온도	12_1_2_4월 일조율	태풍횟수		장마일수	합계강수량	품질		
1	.6078	-1.0861	0.7828	0.8910		0.7946	0.2238	-0.9633		0.3236	-0.9703	2		
-(0.5056	-0.4365	-1.1102	1.3410		-0.1945	0.4505	1.7912		0.5968	-0.5562	2		

[홍로 - 상관분석]

<u>X train</u>									<u>y train</u>	
9월 월합강수량	1월 일조율	12월 일조율	10월 월합강수량		5월 일조율	1월 합계 일사량	태풍횟수	 장마일수	합계강수량	품질
1.7567	-1.1068	0.1083	0.4177	•••	0.6543	0.9436	-0.9379	 0.1772	-0.5493	2
0.0493	-0.7318	0.1531	-0.1820	•••	-1.2951	-0.1977	-0.9379	 -1.0340	-0.1151	1
<u>상관분석을 통해 선별한 10개의 기상 변수</u> <u>추가 변수</u>										

[홍로 - 요인분석]

<u>X train</u>								<u>y train</u>		
1_2월 온도	평균풍속	7월의 기온	5월의 기온		9월 온도	12_1_2_4월 일조율	태풍횟수	 장마일수	합계강수량	품질
-1.45015	0.3391	0.1307	-0.4224		-0.55479	0.3228	0.5734	 -1.0485	-0.5123	2
1.2107	0.0543	0.5438	-1.1452		0.1283	-0.3469	-0.9338	 -0.9564	-0.6990	1

<u>추가 변수</u>

요인분석을 통해 생성한 14개의 기상 변수

04 실험 및 평가

모델 선정 과정 사과 품종 별로 선택 상관분석/요인분석 변수 7가지의 머신러닝 모델 홍로, 후지 각각 총 14가지의 모델

평가지표 소규모 데이터 불균형 데이터 f1 score

다양한 시도 SMOTE 기법 **Grid Search** 최적의 모델 찾기

04 실험 및 평가

데이터의 특성을 기반으로 선택한 14가지의 모델을 돌린 후, 성능에 따라 **후지** 품종에서는 요인분석의 RandomForest, **홍로** 품종에서는 요인분석의 XGBoost를 최종모델로 선정하였다.

후지 상관분석	정확도	f1 score
XGBoost	0.6573	0.6022
RandomForest	0.6573	0.6022
SVM	0.6685	0.6195
GradientBoosting	0.6685	0.6092
AdaBoost	0.6573	0.6022
Voting	0.6573	0.6022
Stacking	0.6573	0.6022

후지 요인분석	정확도	f1 score
XGBoost	0.6685	0.6503
RandomForest	0.6685	0.6503
SVM	0.6685	0.6503
GradientBoosting	0.6404	0.6362
AdaBoost	0.6685	0.6503
Voting	0.6685	0.6503
Stacking	0.6685	0.6503

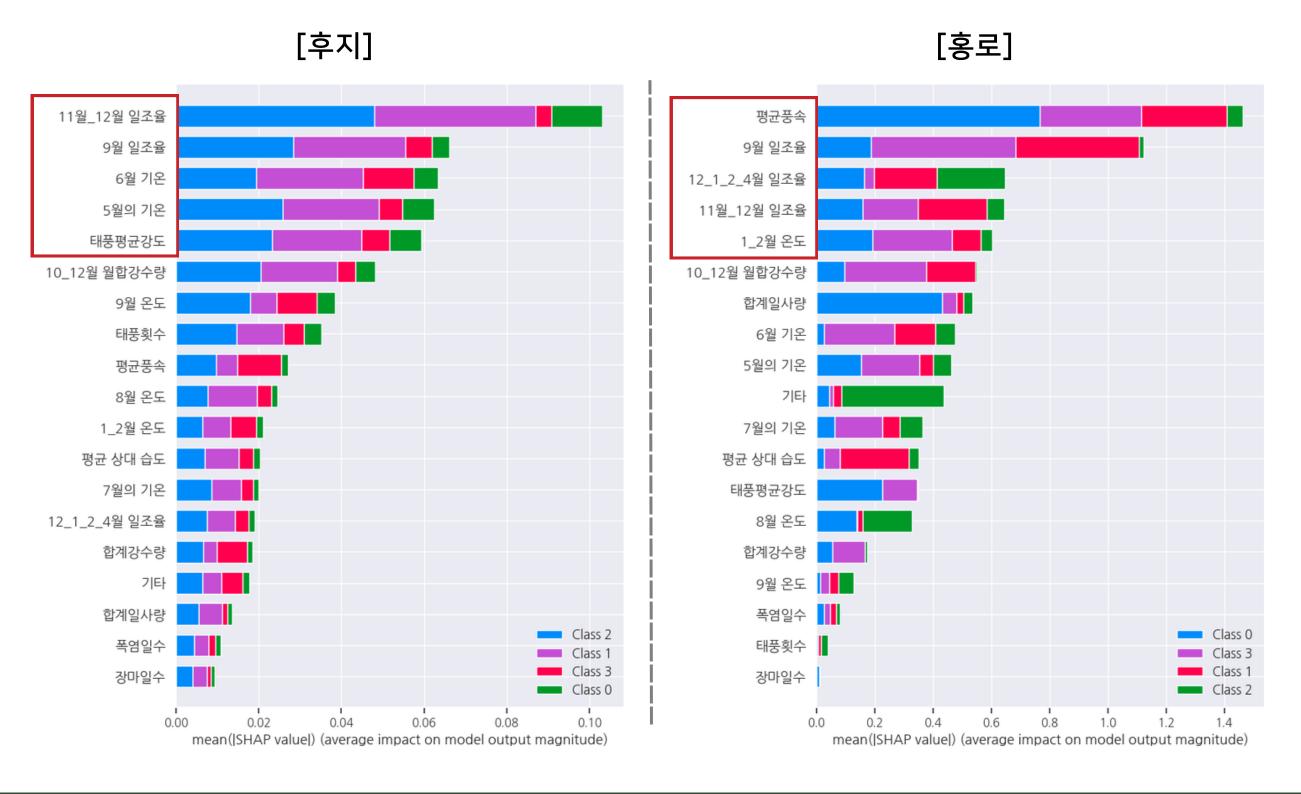
홍로 상관분석	정확도	f1 score
XGBoost	0.6700	0.6468
RandomForest	0.6850	0.6451
SVM	0.6700	0.6438
GradientBoosting	0.6850	0.6451
AdaBoost	0.6900	0.6557
Voting	0.6900	0.6557
Stacking	0.6700	0.6419

홍로 요인분석	정확도	f1 score	
XGBoost	0.7050	0.6787	
RandomForest	0.7100	0.6744	
SVM	0.6800	0.6713	
GradientBoosting	0.7050	0.6787	
AdaBoost	0.6700	0.6303	
Voting	0.6700	0.6303	
Stacking	0.7050	0.6736	

04 실험 및 평가

[SHAP의 summary plot]

모델에서 각 변수가 품질 예측에 얼마나 큰 영향을 미치는지 확인



- SHAP는 Shapley value를 사용하여 피쳐가 모델에 미치는 영향을 측정함.
- SHAP는 피쳐간의 의존성과 음의 영향력을 고려하기 때문에 feature_importances가 보장하지 못하는 넓은 범위의 함정을 피할 수 있음.
- ➡ 따라서, SHAP을 분석 도구로 채택

05 프로그램 개발

프로그램 소개



사과 품질 예측

사용자가 원하는 사과의 품종과 수확지점 및 날짜를 입력받은 후, 해당 지점의 한 해 날씨 데이터를 이용해 사과의 품질을 예측



기상요인 분석

사과의 품질 등급을 결정하는 데 <mark>영향을 끼친</mark> 변수들을 보여주고, 변수별로 끼친 영향을 수치로 나타냄

차별화된 강점



자동화

1년 간의 기상 데이터를 사용자가 직접 입력할 필요 없이, 기상청 API를 이용해 데이터를 자동으로 불러와 사용자의 번거로움을 없앰

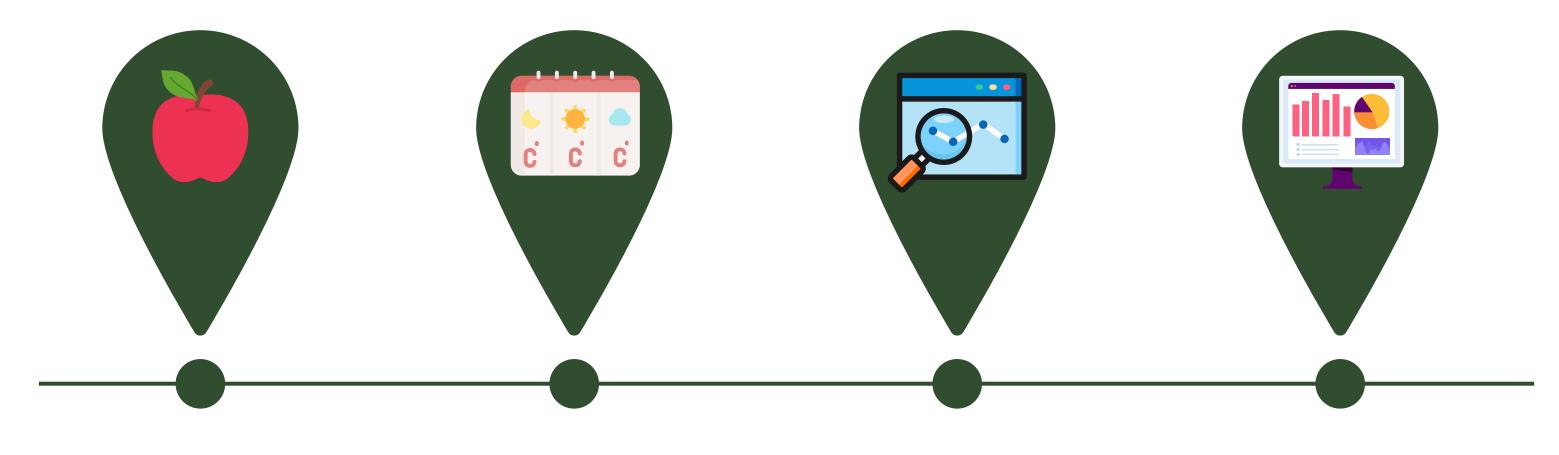


분석력

단순히 예측모델의 결과만 보여주는 것이 아닌 등급별로 사과의 예측 확률을 수치로 나타내고, 품질 예측 과정에서의 변수별 영향력을 나타내 사용자의 해석을 용이하게 하게 함

05 프로그램 개발

[프로그램 작동과정]



01

사과의 품종과 수확지점, 수확날짜를 입력 받는다. 02

수확 지점과 날짜에 따른 기상 데이터를 불러온다. 03

사과의 품종에 따라 미리 최적화된 모델을 불러온 후, 사과의 품질 예측을 수행한다. 04

사과의 예측 품질 등급과 각 등급별 확률, 품질에 영향을 미친 기상 요소를 출력한다.

05 프로그램 개발

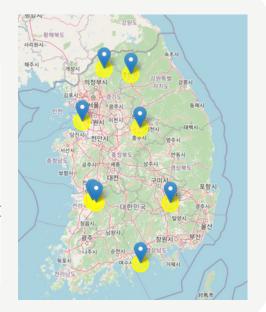
[프로그램 실행 결과]

예측된 품질 등급 및 다른 지점과의 품질 비교

안녕하세요. 사과 품질 예측기입니다.

- ▶ 지점을 입력하세요(화성, 포천, 김제, 완주, 남해, 충주, 춘천, 군위): 춘천
- ▶ 사과 종류를 입력하세요(후지, 홍로): 후지
- ▶ 수확 날짜를 입력하세요(yyyy/MM/DD 형식): 2018/10/10
- ▶ 같은 날 수확한 다른 지점의 사과 품질도 함께 출력을 원하십니까?(5분 이상의 시간이 소요됩니다.(0,X)): (

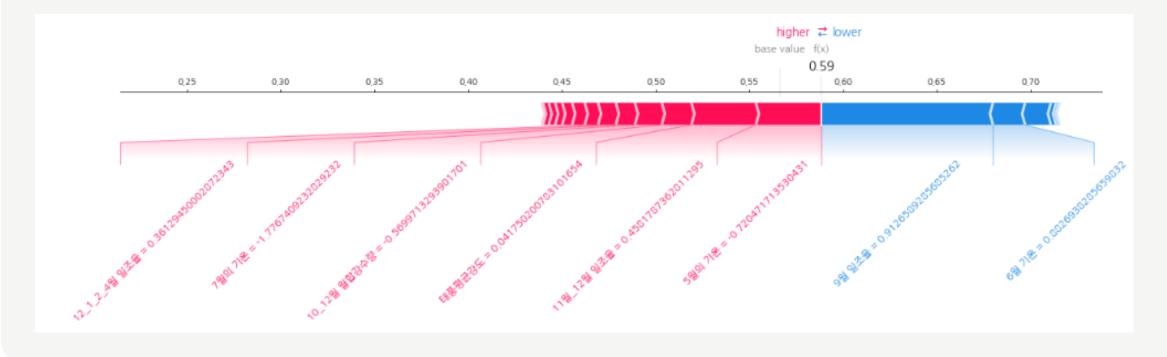
요청하신 춘천에서 생산된 후지 사과의 품질은 [3] 등급 입니다.



품질 등급별 예측 확률



품질에 영향을 미친 기상요소





06 향후 계획 및 기대효과

01차별화된 강점

날씨 데이터만으로 <mark>사과의 품질을 예측</mark>할 수 있고, 어떤 기상요소가 사과 품질에 중요한지 파악할 수 있다.

요인분석을 진행해 변수 간의 복잡한 상호작용을 이해하고 변수 간의 연관성을 발견할 수 있으며, 이를 이용하여 변수를 축약할 수 있다.

API를 이용해 프로그램 사용자의 추가 데이터 입력 없이 자동으로 사과 품질예측이 가능하다.

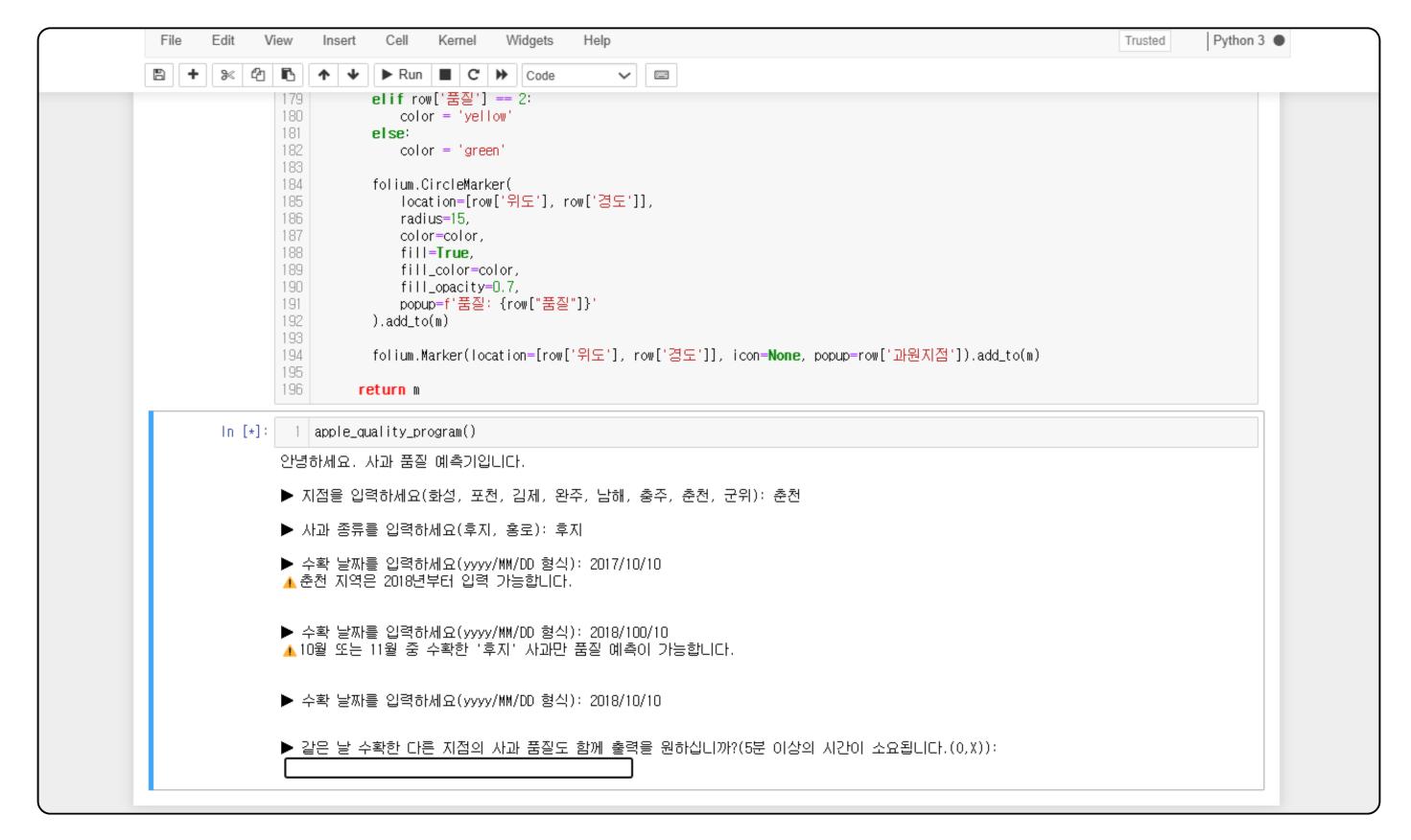
02 기대효과

사과의 품질과 기후 요소들 간의 상관관계를 파악한다면, 농가들은 더욱 효율적인 생산 방법을 도입하고 농작물 관리에 있어 더 나은 결정을 내릴 수 있다.

소비자들은 예측 가능한 품질의 사과를 구매함으로써 <mark>만족도가 향상</mark>되며, 이는 사과 시장에 긍정적인 영향을 끼칠 것이다.

스마트팜 건설, 품종 개량 등 현실적인 문제에 알고리즘을 적용할 수 있다.

07 시연



<u>링크: https://www.youtube.com/watch?v=bGH5zOxep5A</u>

THANK YOU

지금까지 아삭팀의 발표를 들어주셔서 감사합니다.