

파이썬 Streamlit 라이브러리를 활용한 IRIS 데이터 머신러닝 대시보드 개발

임여명 수강생

2022년 12월 8일

메뉴

HOME ▼

강의명	(산대특)_공공데이터 활용 빅데이터 분석 및 시각화 전문가 과정 육성
교과목명	기초문법
프로젝트 주제	파이썬 Streamlit 라이브러리를 활용한 IRIS 데이터 머신러닝 대시보드 개발
프로젝트 마감일	2022년 12월 8일
수강생명	임여명

Home

IRIS 예측 모델 개발

- IRIS 데이터를 활용하여 간단한 EDA 및 예측 모델을 구현한다.

데이터 출처

<https://www.kaggle.com/datasets/saurabh00007/iris.csv>

데이터셋 정보

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1000	3.5000	1.4000	0.2000	Iris-setosa
1	4.9000	3.0000	1.4000	0.2000	Iris-setosa
2	4.7000	3.2000	1.3000	0.2000	Iris-setosa
3	4.6000	3.1000	1.5000	0.2000	Iris-setosa
4	5.0000	3.6000	1.4000	0.2000	Iris-setosa
5	5.4000	3.9000	1.7000	0.4000	Iris-setosa
6	4.6000	3.4000	1.4000	0.3000	Iris-setosa
7	5.0000	3.4000	1.5000	0.2000	Iris-setosa
8	4.4000	2.9000	1.4000	0.2000	Iris-setosa
9	4.9000	3.1000	1.5000	0.1000	Iris-setosa

기술 통계량

	count	mean	std	min	25%	50%	75%	max
sepal_length	150.0000	5.8433	0.8281	4.3000	5.1000	5.8000	6.4000	7.9000
sepal_width	150.0000	3.0573	0.4359	2.0000	2.8000	3.0000	3.3000	4.4000
petal_length	150.0000	3.7580	1.7653	1.0000	1.6000	4.3500	5.1000	6.9000
petal_width	150.0000	1.1993	0.7622	0.1000	0.3000	1.3000	1.8000	2.5000

타겟분포

	species
Iris-setosa	50
Iris-versicolor	50
Iris-virginica	50

데이터 타입

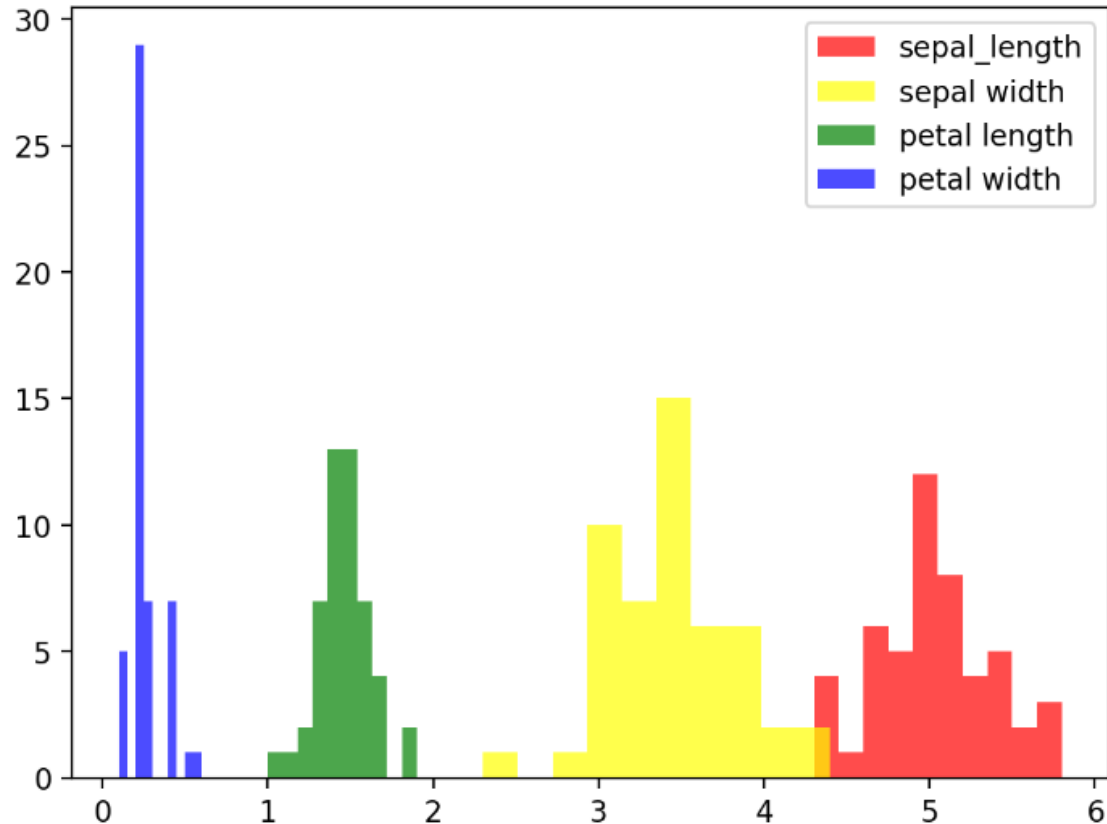
↑	sepal_length	sepal_width	petal_length	petal_width	species
구분	float64	float64	float64	float64	object

기술통계량 서브메뉴에서 Iris.csv의 데이터셋 정보, 기술통계량, 데이터 타입, 타겟분포 등을 조회할 수 있습니다.

탐색적 자료 분석 - 2

히스토그램 산점도 박스플롯

붓꽃의 세 종류인 Setosa, versicolor, virginic 의 꽃받침 길이(Sepal Length), 꽃받침 폭(Sepal Width), 꽃잎 길이(Petal Length),꽃잎 폭(Petal Width)을 히스토그램으로 나타낸 것입니다.

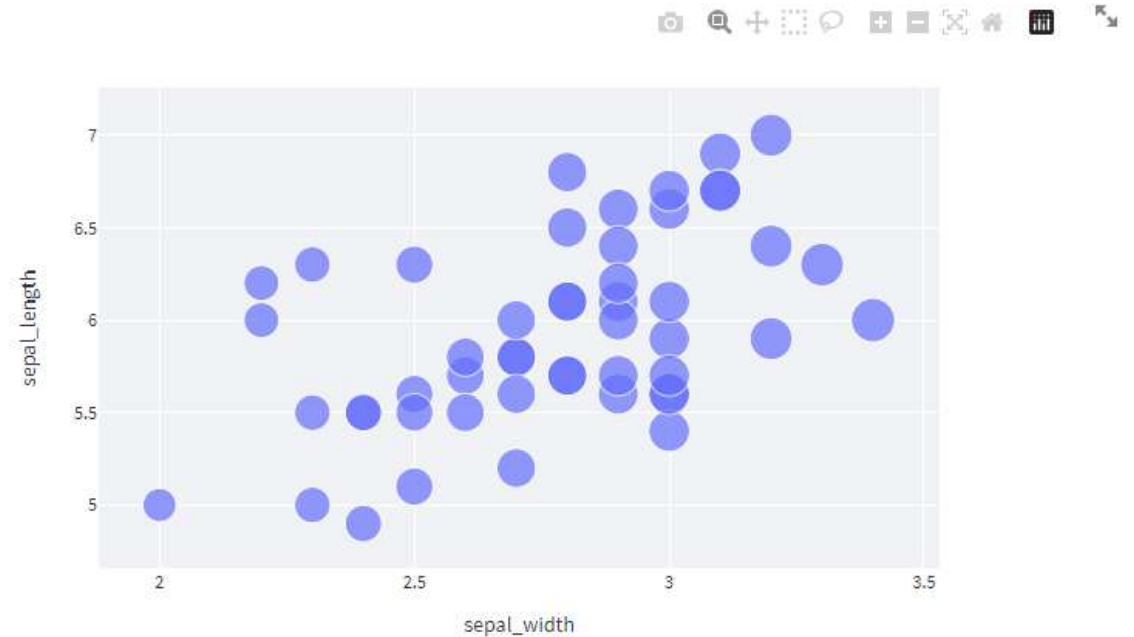


종 선택

Iris-setosa

히스토그램 산점도 박스플롯

붓꽃의 세 종류인 Setosa, versicolor, virginic 의 꽃받침 길이(Sepal Length), 꽃받침 폭(Sepal Width)을 산점도로 나타낸 것입니다.



종 선택

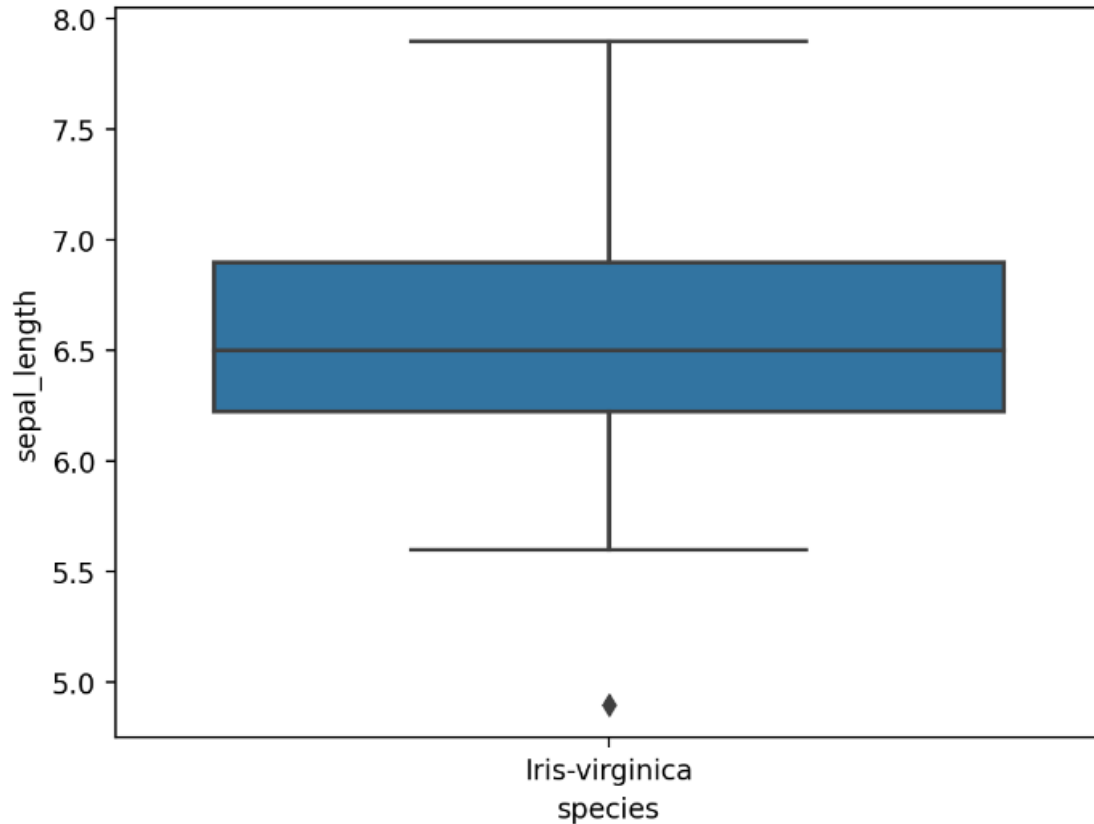
Iris-versicolor

종 선택: Iris-versicolor

그래프 서브메뉴에서는 붓꽃의 각 종류와 히스토그램, 산점도, 박스플롯의 세가지 시각화 그래프 중 하나를 선택할 수 있습니다.

히스토그램 산점도 **박스플롯**

붓꽃의 세 종류인 Setosa, versicolor, virginic 의 꽃받침 길이(Sepal Length), 종(species)을 박스플롯으로 나타낸 것입니다.

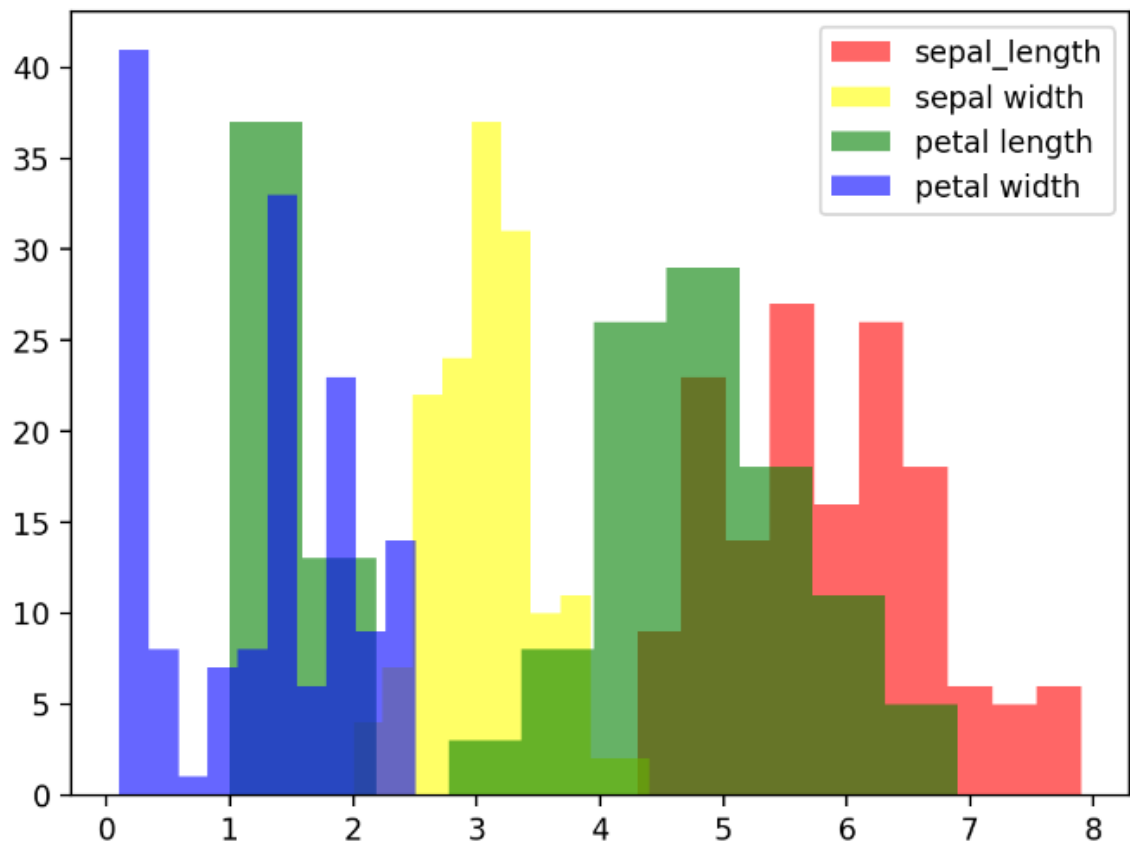


종 선택

Iris-virginica

그래프 서브메뉴에서는 붓꽃의 각 종류와 히스토그램, 산점도, 박스플롯의 세가지 시각화 그래프 중 하나를 선택할 수 있습니다.

히스토그램 종합

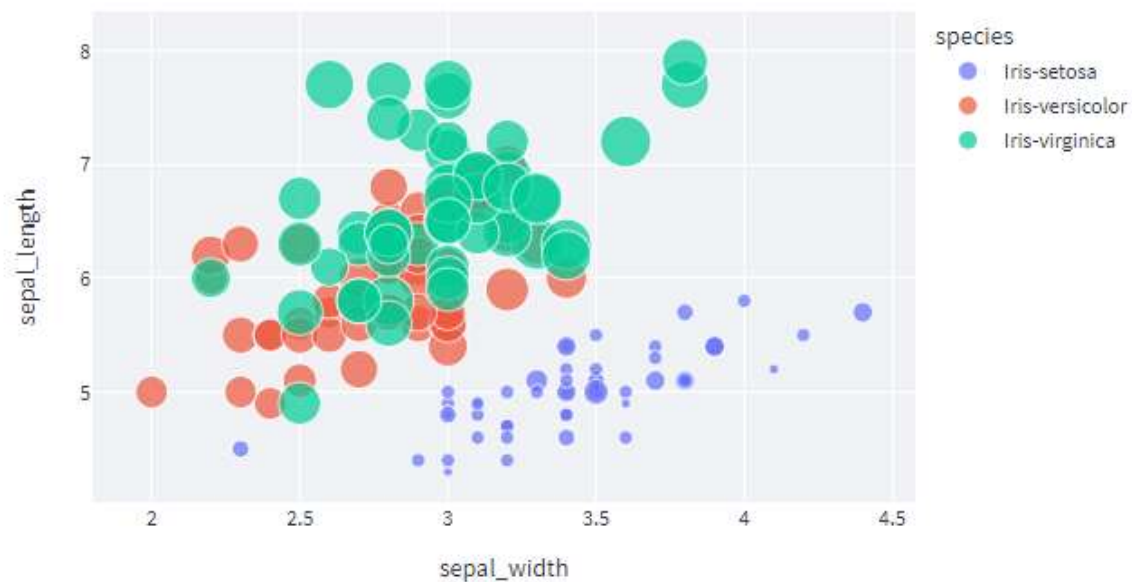


산점도 종합

박스플롯 종합

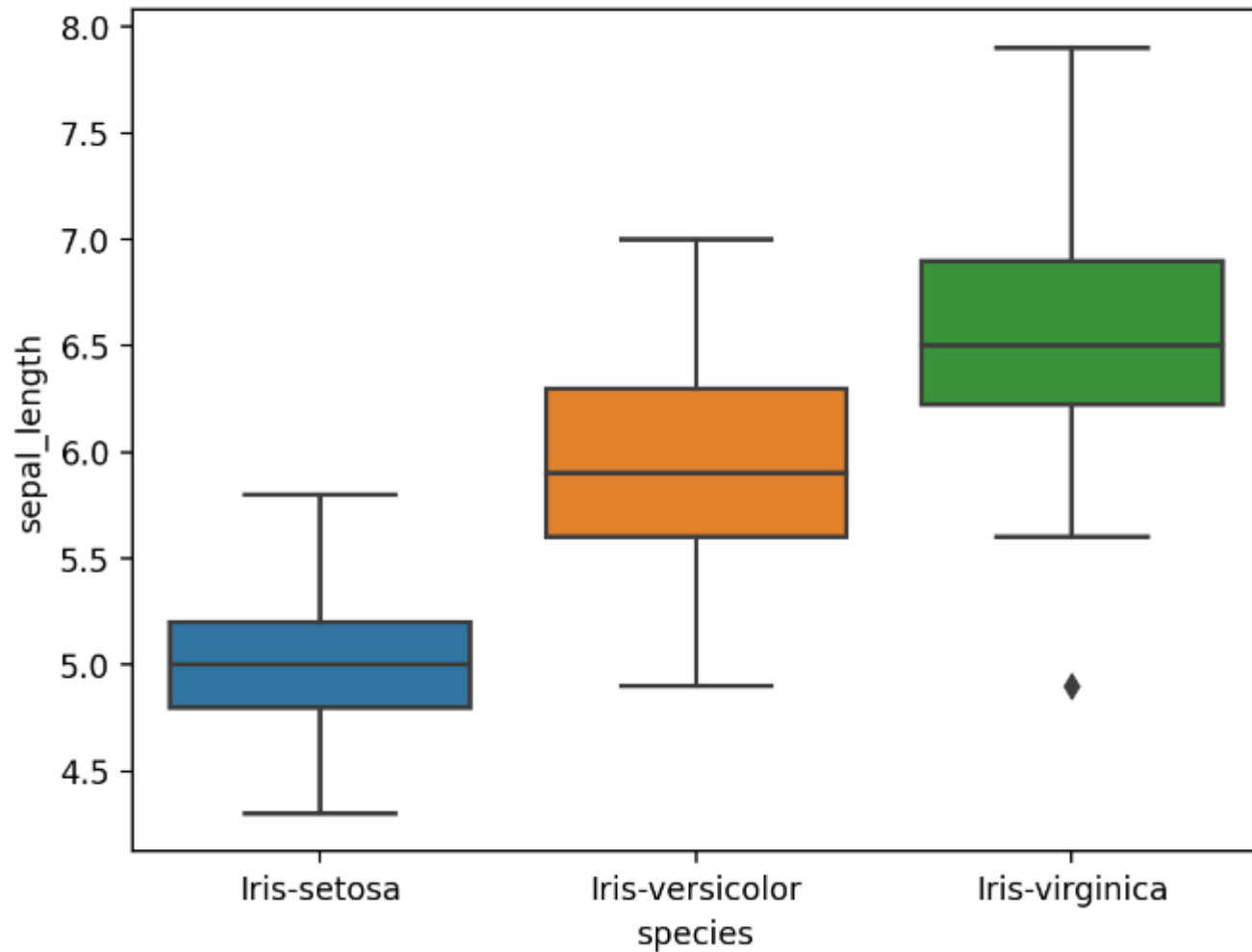
산점도 종합

IRIS 산점도



세 종류의 붓꽃 데이터를 하나의 그래프로 시각화한 종합 그래프를 확인할 수 있습니다.

박스플롯 종합

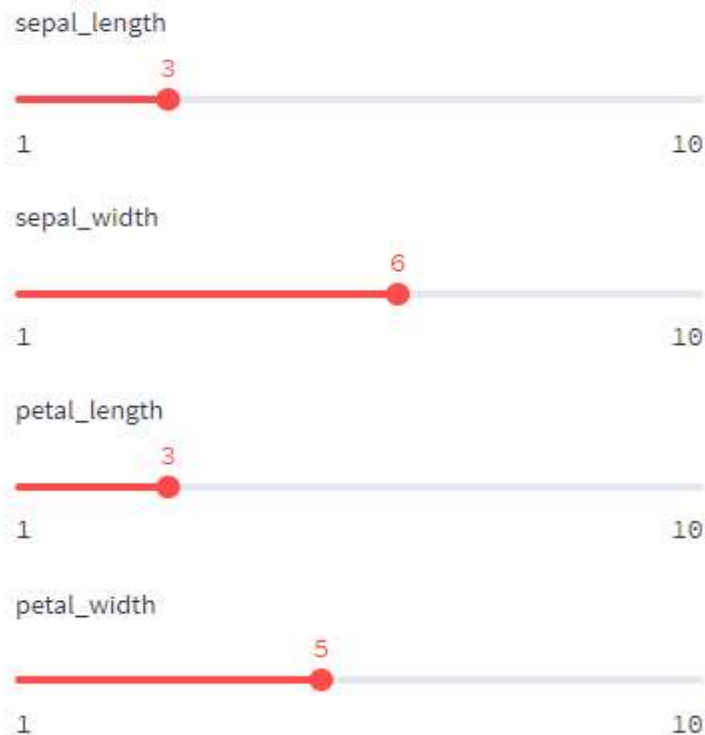


세 종류의 붓꽃 데이터를 하나의 그래프로 시각화한 종합 그래프를 확인할 수 있습니다.

속성 정보:

1. 꽃받침 길이(cm) - sepal length
2. 꽃받침 폭(cm) - sepal width
3. 꽃잎 길이(cm) - petal length
4. 꽃잎 너비(cm) - petal width
5. 종류:
 - Iris Setosa
 - Iris Versicolour
 - Iris Virginica

입력값을 조정하세요.



예측 결과를 확인해주세요.

	0	1	2
0	0.1037	0.0004	0.8958

Virginica 종입니다.

Setosa 확률: 10.37 %

Versicolor 확률: 0.04 %

Virginica 확률: 89.58 %



sepal_length, sepal_width, petal_length, petal_width 각각의 길이를 설정하면 입력값에 따라 확률을 표기해주고
붓꽃의 종류가 어떤 것인지 예측합니다.

강의명	(산대특)_공공데이터 활용 빅데이터 분석 및 시각화 전문가 과정 육성
교과목명	기초문법
프로젝트 주제	파이썬 Streamlit 라이브러리를 활용한 IRIS 데이터 머신러닝 대시보드 개발
프로젝트 마감일	2022년 12월 8일
수강생명	임여명

About

임여명 github : <https://github.com/ryeomyoung2/streamlit-app-human>