

# Lecture 21

Keras를 이용한 딥러닝 코드 분석

다중 클래스 분류(Multi-class Classification) 방식을 이용한  
아이리스 품종 예측

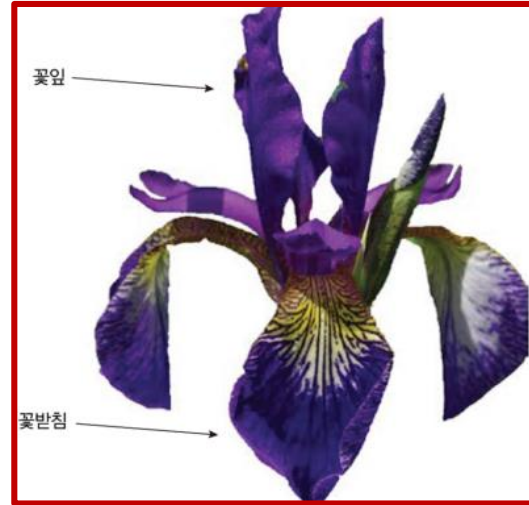
## 1. 아이리스 품종 예측을 위한 데이터셋

## 2. 다중 클래스 분류 (Multi-class Classification)

## 3. Softmax 이해

## 4. pairplot() 함수를 이용한 데이터 분석

## 5. Softmax를 이용한 딥러닝 설계 및 구현



버지니카



Iris-virginica

세토사



Iris-setosa

버시컬러



Iris-versicolor

- 아이리스는 그 꽃봉오리가 마치 먹물을 머금은 붓과 같다하여 우리나라에서 '붓꽃'이라고 불리는 아름다운 꽃
- 아이리스는 **꽃받침**과 **꽃잎**의 **모양**과 **길이**에 따라 여러 가지 품종으로 나뉨
- 과연 딥러닝을 사용하여 이들을 구별해 낼 수 있을까?

## 1. 아이리스 품종 예측을 위한 데이터셋

## 2. 다중 클래스 분류 (Multi-class Classification)

## 3. Softmax 이해

## 4. pairplot() 함수를 이용한 데이터 분석

## 5. Softmax를 이용한 딥러닝 설계 및 구현

아이리스 품종 예측을 위한 데이터 → **iris.csv**

### ■ 데이터의 구조

데이터의 구조

속성					클래스	
	sepal_length	sepal_width	petal_length	petal_width	species	
샘플	0	5.1	3.5	1.4	0.2	Iris-setosa
	1	4.9	3	1.4	0.2	Iris-setosa
	2	4.7	3.2	1.3	0.2	Iris-setosa
	3	4.6	3.1	1.5	0.2	Iris-setosa
	4	5	3.6	1.4	0.2	Iris-setosa

- 샘플 수: 150
- 속성 수: 4
  - 정보 1: 꽃받침 길이 (sepal length, 단위: cm)
  - 정보 2: 꽃받침 넓이 (sepal width, 단위: cm)
  - 정보 3: 꽃잎 길이 (petal length, 단위: cm)
  - 정보 4: 꽃잎 넓이 (petal width, 단위: cm)
- 클래스: Iris-setosa, Iris-versicolor, Iris-virginica  
세토사, 버시컬러, 버지니카

참(1)과 거짓(0)으로 해결하는 것이 아니라  
여러 개 중에 어떤 것이 답인지를 예측하는 문제  
→ **다중 클래스 분류(Multi-class Classification)**

1. 아이리스 품종 예측을 위한 데이터셋

## 2. 다중 클래스 분류 (Multi-class Classification)

3. Softmax 이해

4. pairplot() 함수를 이용한 데이터 분석

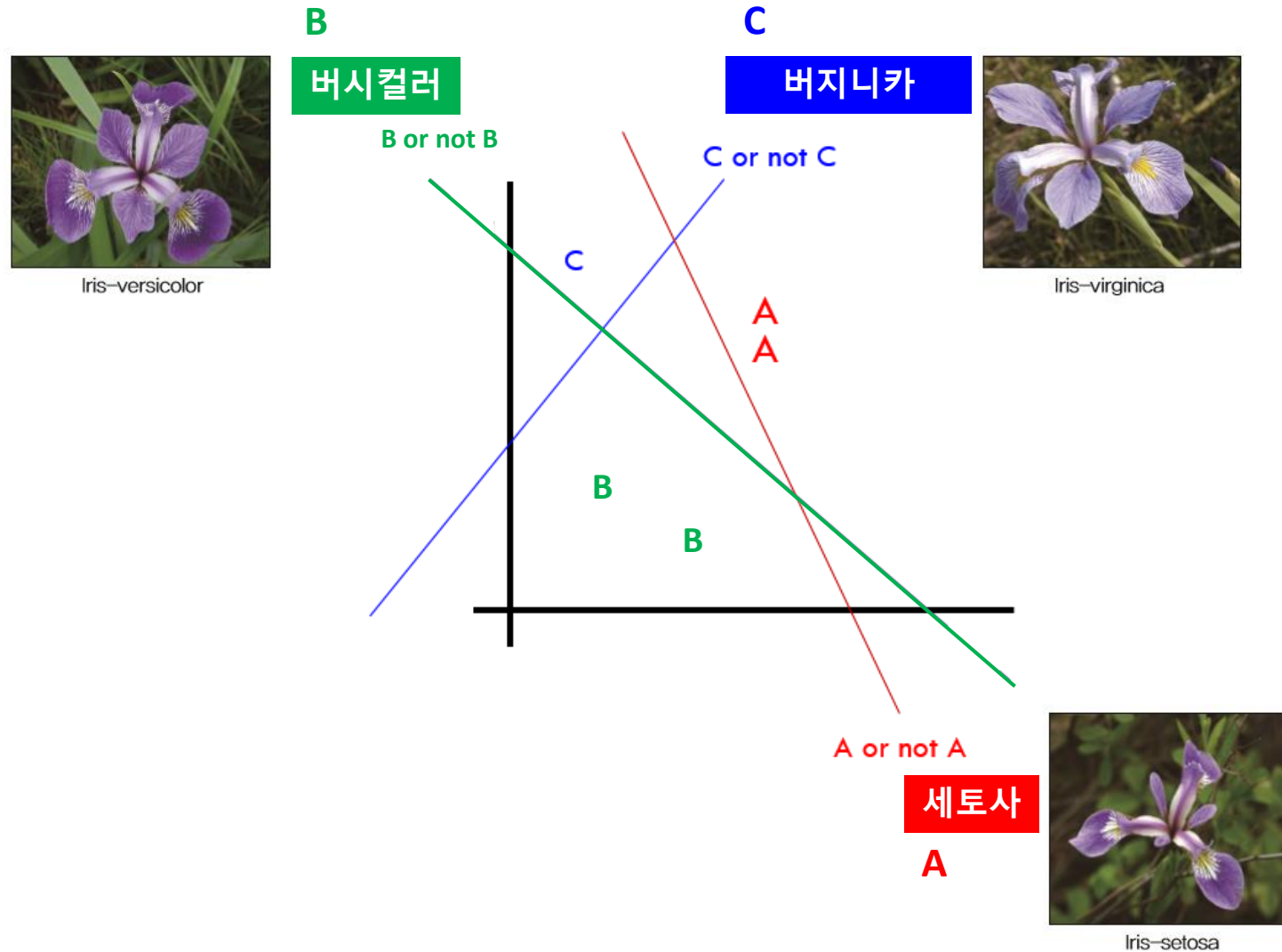
5. Softmax를 이용한 딥러닝 설계 및 구현

### ■ 다중 클래스 분류(multi-class classification)

- ✓ 여러 개의 답 중 하나를 고르는 **분류(classification)** 문제
- ✓ 앞서 다루었던 것과 중요한 차이는 바로 **클래스**가 2개가 아니라 **3개**임
- ✓ 참(1)과 거짓(0)으로 해결하는 것이 아니라, 여러개 중에 어떤 것이 답인지를 예측하는 문제
- ✓ **다중 클래스 분류** 문제는 둘 중에 하나를 고르는 **이항 분류(binary classification)**와는 접근 방식이 조금 다름

1. 아이리스 품종 예측을 위한 데이터셋
2. 다중 클래스 분류  
(Multi-class Classification)
3. **Softmax** 이해
4. pairplot() 함수를 이용한 데이터 분석
5. Softmax를 이용한 딥러닝 설계 및 구현

## 이항 분류(binary classification)



1. 아이리스 품종 예측을 위한 데이터셋

2. 다중 클래스 분류  
(Multi-class Classification)

### 3. Softmax 이해

4. pairplot() 함수를 이용한 데이터 분석

5. Softmax를 이용한 딥러닝 설계 및 구현

## 이항 분류(binary classification)



Iris-setosa

세토사

$[W_{A1} \ W_{A2} \ W_{A3}]$

$\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix}$

$$= [W_{A1}x1 + W_{A2}x2 + W_{A3}x3]$$

$x \rightarrow$

$W$

$z \rightarrow$

$s$

$\hat{Y}$



Iris-versicolor

버시컬러

$[W_{B1} \ W_{B2} \ W_{B3}]$

$\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix}$

$$= [W_{B1}x1 + W_{B2}x2 + W_{B3}x3]$$

$x \rightarrow$

$W$

$z \rightarrow$

$s$

$\hat{Y}$



Iris-virginica

버지니카

$[W_{C1} \ W_{C2} \ W_{C3}]$

$\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix}$

$$= [W_{C1}x1 + W_{C2}x2 + W_{C3}x3]$$

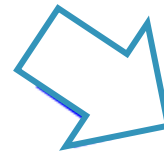
$x \rightarrow$

$W$

$z \rightarrow$

$s$

$\hat{Y}$



세토사

$[W_{A1} \ W_{A2} \ W_{A3}]$

버시컬러

$[W_{B1} \ W_{B2} \ W_{B3}]$

버지니카

$[W_{C1} \ W_{C2} \ W_{C3}]$

$\begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix}$

$=$

## 이항 분류(binary classification)

1. 아이리스 품종 예측을 위한 데이터셋

2. 다중 클래스 분류  
(Multi-class Classification)

### 3. Softmax 이해

4. pairplot() 함수를 이용한 데이터 분석

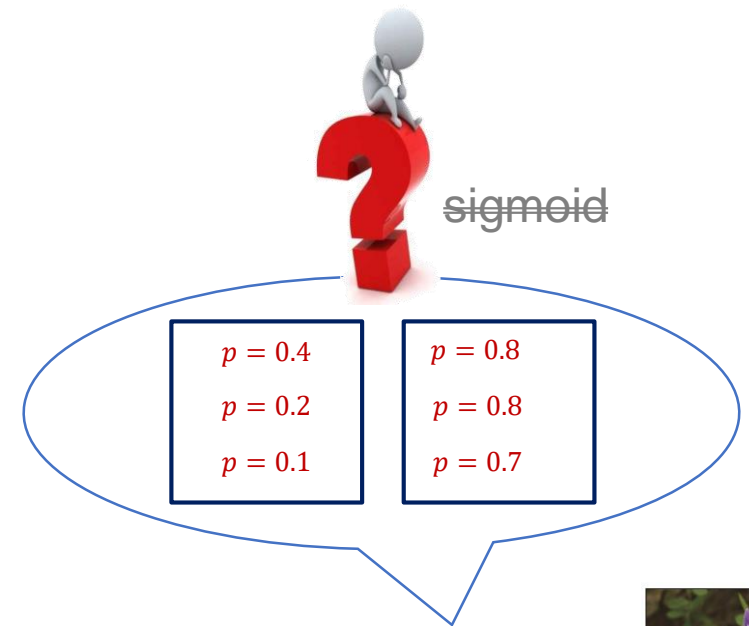
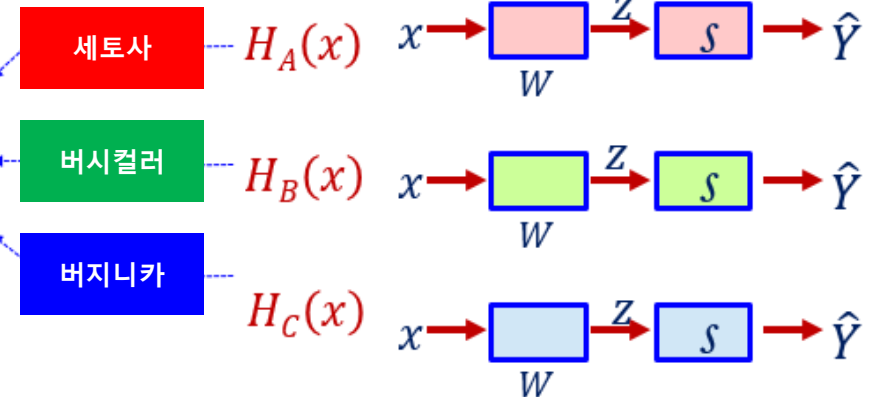
5. Softmax를 이용한 딥러닝 설계 및 구현

$$\begin{bmatrix} \text{세토사} \\ \text{버시컬러} \\ \text{버지니카} \end{bmatrix} \begin{bmatrix} W_{A1} & W_{A2} & W_{A3} \\ W_{B1} & W_{B2} & W_{B3} \\ W_{C1} & W_{C2} & W_{C3} \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} =$$



$$\begin{bmatrix} w_{A1}x1 + w_{A2}x2 + w_{A3}x3 \\ w_{B1}x1 + w_{B2}x2 + w_{B3}x3 \\ w_{C1}x1 + w_{C2}x2 + w_{C3}x3 \end{bmatrix} =$$

$$\begin{bmatrix} \hat{Y}_A \\ \hat{Y}_B \\ \hat{Y}_C \end{bmatrix}$$



1. 아이리스 품종 예측을 위한 데이터셋
2. 다중 클래스 분류  
(Multi-class Classification)
3. Softmax 이해
4. pairplot() 함수를 이용한 데이터 분석
5. Softmax를 이용한 딥러닝 설계 및 구현

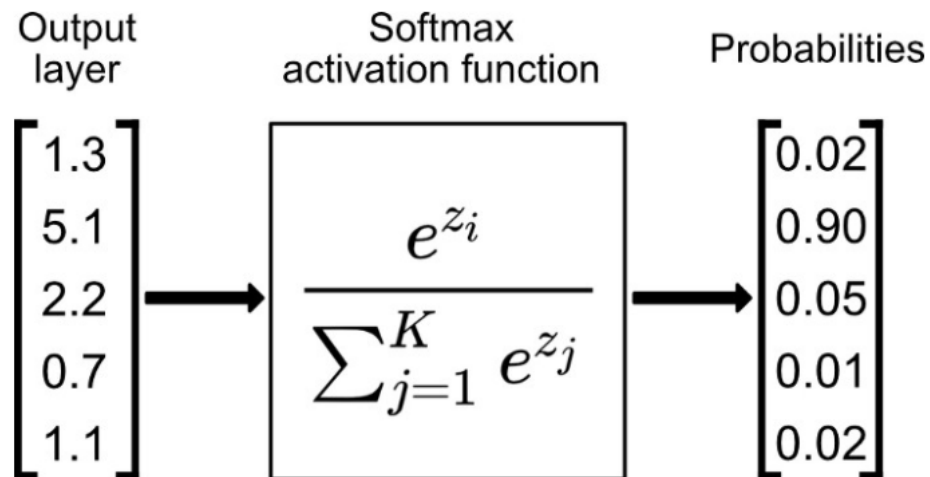
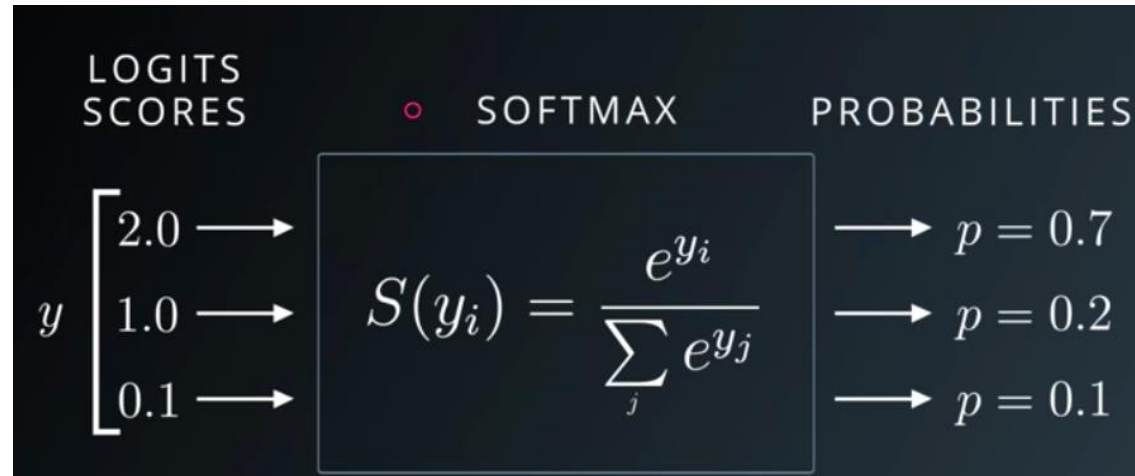
## 다중 클래스 분류(multi-class classification)

if) 3개의 클래스

$$wx + b = y$$

~~sigmoid~~

if) 5개의 클래스





1. 아이리스 품종 예측을 위한 데이터셋

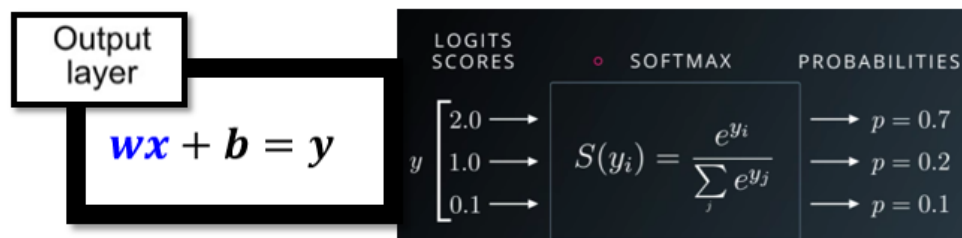
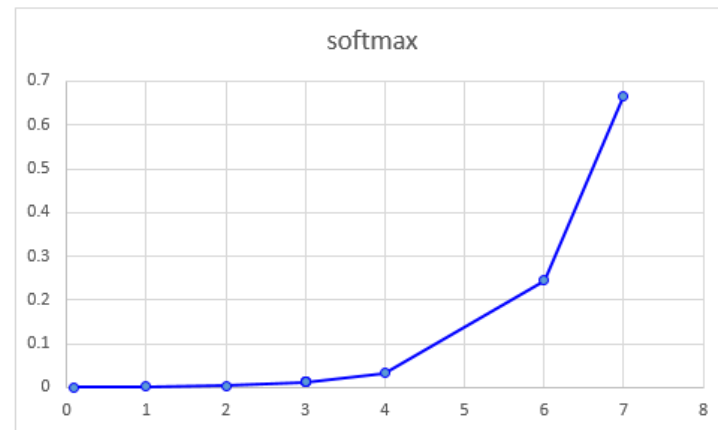
2. 다중 클래스 분류  
(Multi-class Classification)

### 3. Softmax 이해

4. pairplot() 함수를 이용한 데이터 분석

5. Softmax를 이용한 딥러닝 설계 및 구현

	$y$	$e^{y_i}$	$S(y_i)$ SOFTMAX
자연상수	0.1	1.10517092	0.000671792
2.718281828	1	2.71828183	0.001652341
	2	7.3890561	0.004491528
	3	20.0855369	0.012209239
	3	20.0855369	0.012209239
	3	20.0855369	0.012209239
	3	20.0855369	0.012209239
	4	54.59815	0.033188154
	6	403.428793	0.24522913
	7	1096.63316	0.666601889
sum		$\sum_j e^{y_j}$ 1645.10959	1



- Softmax는 Normalized Exponential Function이라고도 불림
- $x$ 가 최대값 근처에서 거의 대부분의 확률을 가져간다.  
위 그래프에서는 최대값인 7 이 선택될 확률이 가장 높아 진다.  
7 이라는 숫자는 앞 단계에서 계산해 놓은  $w\mathbf{x}$  에 해당하는 값이다.

1. 아이리스 품종 예측을 위한  
데이터셋

2. 다중 클래스 분류  
(Multi-class Classification)

### 3. Softmax 이해

4. pairplot() 함수를 이용한  
데이터 분석

5. Softmax를 이용한 딥러닝  
설계 및 구현

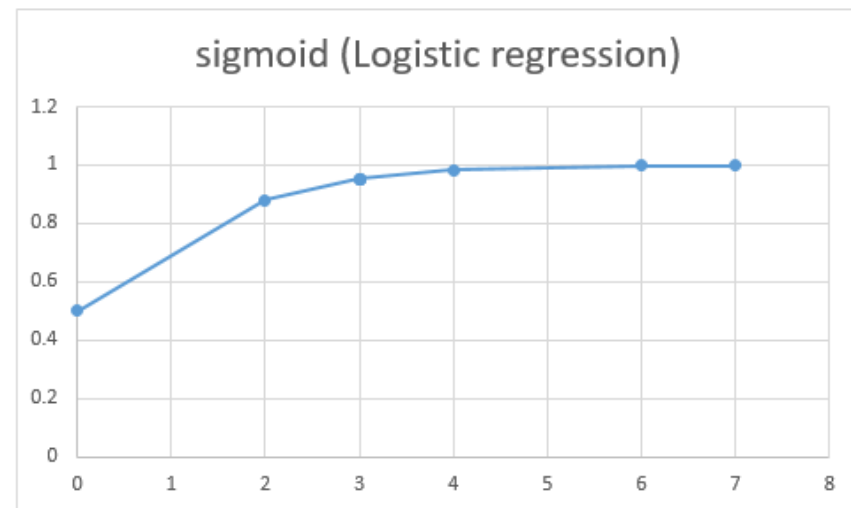
	$y$	sigmoid
자연상수	0	0.5
<b>2.718281828</b>	2	0.880797078
	3	0.952574127
	3	0.952574127
	3	0.952574127
	3	0.952574127
	4	0.98201379
	6	0.997527377
	7	0.999088949

Output  
layer

$$wx + b = y$$

시그모이드 함수를 나타내는 방정식

$$y = \frac{1}{1 + e^{-(ax+b)}}$$



1. 아이리스 품종 예측을 위한 데이터셋

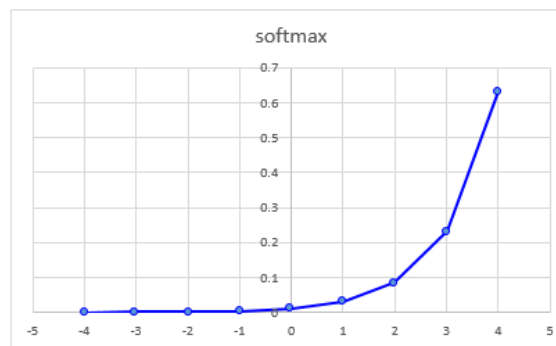
2. 다중 클래스 분류  
(Multi-class Classification)

### 3. Softmax 이해

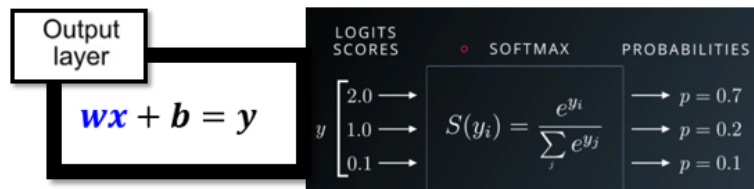
4. pairplot() 함수를 이용한 데이터 분석

5. Softmax를 이용한 딥러닝 설계 및 구현

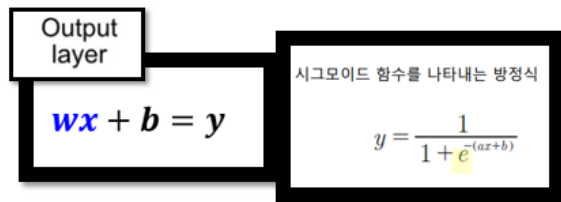
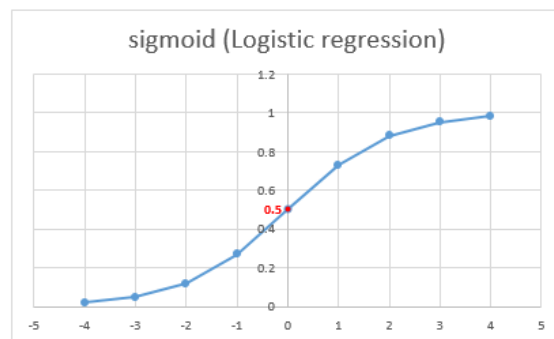
	$y$	$e^{y_i}$	$S(y_i)$ SOFTMAX
자연상수	-4	0.018315639	0.000212124
2.718281828	-3	0.049787068	0.000576613
	-2	0.135335283	0.001567396
	-1	0.367879441	0.004260624
	0	1	0.011581577
	1	2.718281828	0.031481991
	2	7.389056099	0.085576923
	3	20.08553692	0.232622194
	4	54.59815003	0.632332683
sum		$\sum e^{y_j}$ 86.34402668	1



- x축을 -4 에서 4 까지의 범위를 입력으로 하면 로지스틱 회귀에서는 "S"형태의 시그모이드 결과가 출력된다. 이 경우에도 Softmax는 x가 최대값 근처에서 거의 대부분의 확률을 가져간다.

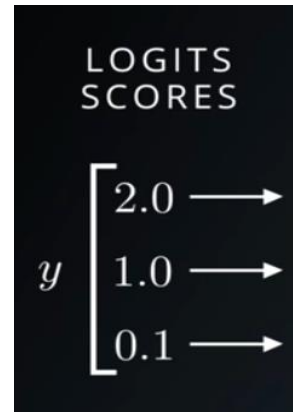
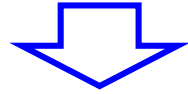


	$y$	sigmoid
자연상수	-4	0.01798621
2.718281828	-3	0.047425873
	-2	0.119202922
	-1	0.268941421
	0	0.5
	1	0.731058579
	2	0.880797078
	3	0.952574127
	4	0.98201379



1. 아이리스 품종 예측을 위한 데이터셋
2. 다중 클래스 분류  
(Multi-class Classification)
3. Softmax 이해
4. pairplot() 함수를 이용한 데이터 분석
5. Softmax를 이용한 딥러닝 설계 및 구현

$$wx + b = \hat{Y}$$



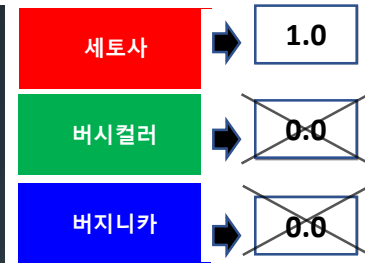
SOFTMAX

$$S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}}$$

Probabilities

→	$p = 0.7$
→	$p = 0.2$
→	$p = 0.1$

1.0



Iris-setosa



Iris-versicolor



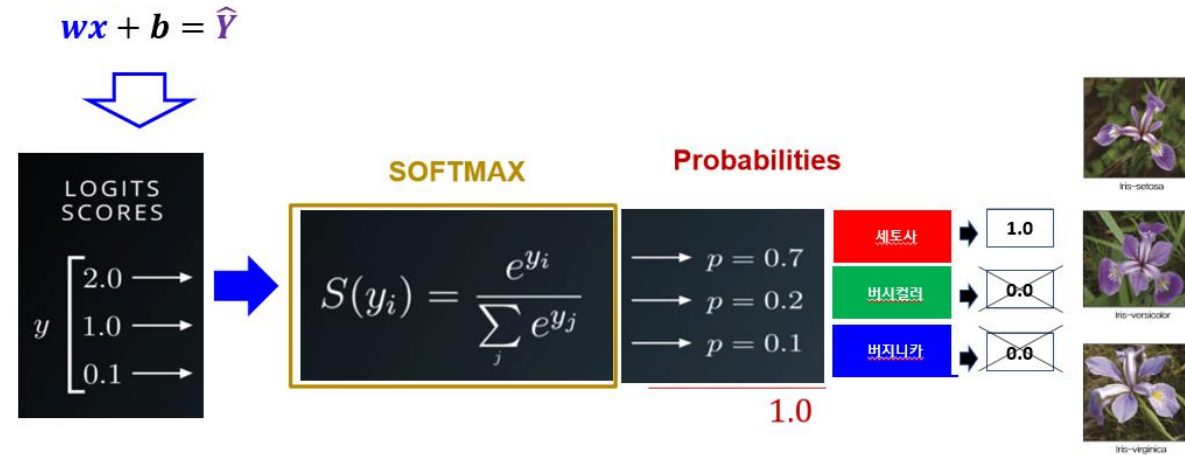
Iris-virginica

- 여러 개의 연산 결과를 정규화하여 모든 클래스의 확률값 합을 1로 만들자는 간단한 아이디어
- 정규화 함수에 자연상수를 한 번 씌워주는데, cost function의 미분값을 convex하게 만들어주기 위한 것
- Logistic Regression과 Cross-entropy에 대해 제대로 이해했다면 연관지어서 생각하면 됨

1. 아이리스 품종 예측을 위한 데이터셋
2. 다중 클래스 분류  
(Multi-class Classification)

### 3. Softmax 이해

4. pairplot() 함수를 이용한 데이터 분석
5. Softmax를 이용한 딥러닝 설계 및 구현



<https://scikit-learn.org/stable/>

from sklearn.preprocessing import LabelEncoder 으로 부터 제공됨

LabelEncoder() : 텍스트 형식의 클래스를 숫자로 변환

사이킷런(scikit-learn)을  
이용한 코딩

species	index	One-Hot		
세토사	0	1	0	0
버시컬러	1	0	1	0
버지니카	2	0	0	1

tf.keras.utils.to\_categorical(index) :

“ONE-HOT” Encoding 으로 변환

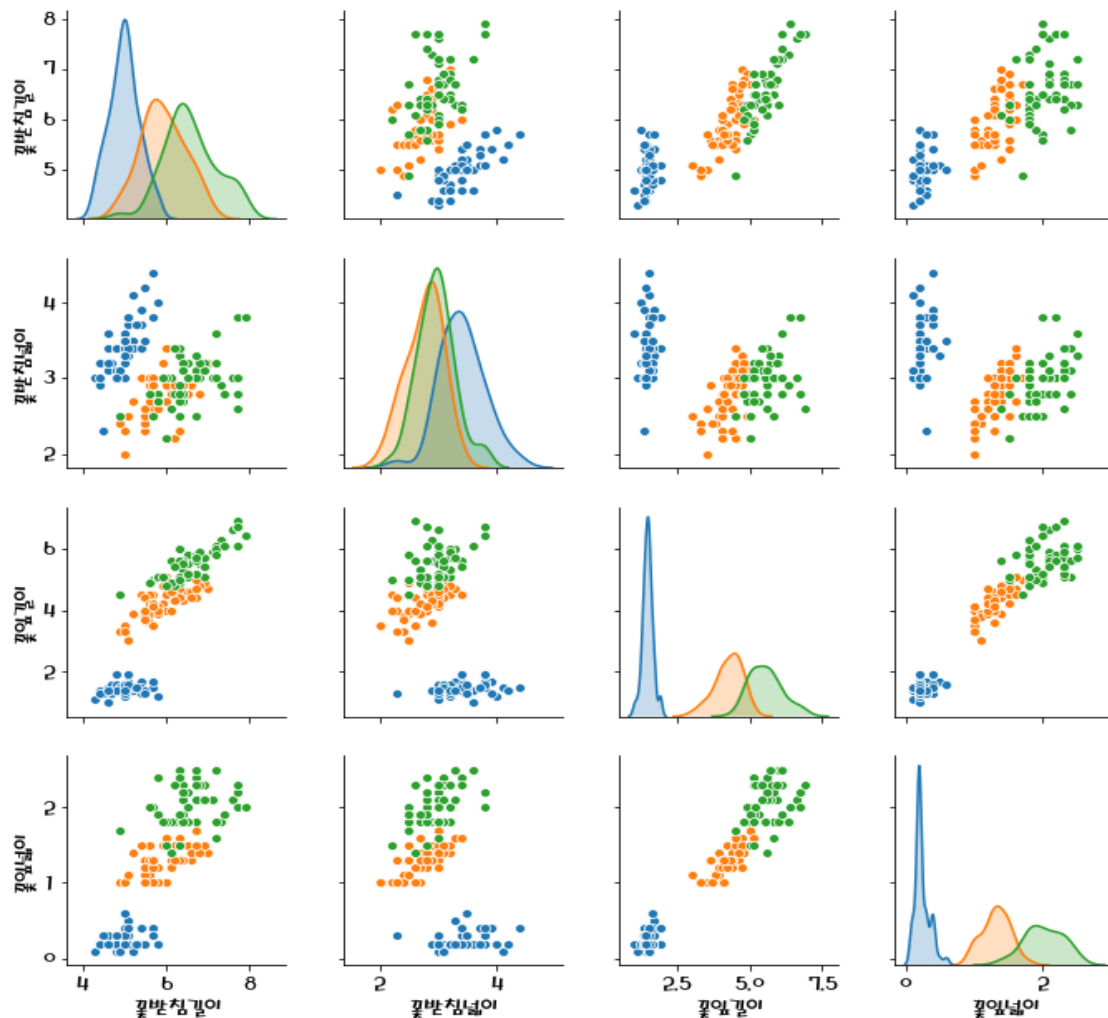
Keras를  
이용한 코딩

1. 아이리스 품종 예측을 위한 데이터셋
2. 다중 클래스 분류  
(Multi-class Classification)
3. Softmax 이해
4. **pairplot() 함수를 이용한 데이터 분석**
5. Softmax를 이용한 딥러닝 설계 및 구현

seaborn에서 제공하는 **pairplot()** 함수로 데이터 한번에 보기

<https://seaborn.pydata.org/>

[https://seaborn.pydata.org/examples/scatterplot\\_matrix.html](https://seaborn.pydata.org/examples/scatterplot_matrix.html)



- 사진으로 볼 때는 비슷해 보이던 **꽃잎**과 **꽃받침**의 길이와 넓이가 품종별로 차이가 있음을 알 수 있음
- 속성별로 어떤 연관이 있는지를 보여주는 **상관도 그래프**를 통해 프로젝트의 계획을 세우고 프로그램 전략을 세울 수 있음

종류

- *Iris-setosa*
- *Iris-versicolor*
- *Iris-virginica*

1. 아이리스 품종 예측을 위한 데이터셋
2. 다중 클래스 분류  
(Multi-class Classification)
3. Softmax 이해
4. pairplot() 함수를 이용한 데이터 분석
5. **Softmax를 이용한 딥러닝 설계 및 구현**

- 딥러닝 설계 및 구현

아이리스 품종 예측을 위한 데이터 → [iris.csv](#)



21\_(15 page) 강의용\_ Iris.ipynb

1. 아이리스 품종 예측을 위한 데이터셋
2. 다중 클래스 분류  
(Multi-class Classification)
3. Softmax 이해
4. pairplot() 함수를 이용한 데이터 분석
5. Softmax를 이용한 딥러닝 설계 및 구현

- 딥러닝 설계 및 구현

아이리스 품종 예측을 위한 데이터 → [iris.csv](#)



**validation\_data를 사용하여 학습 시, 학습 상황을 확인하는 방법**

**21\_(16 page) 강의용\_Iris (Validation).ipynb**



1. 아이리스 품종 예측을 위한 데이터셋
2. 다중 클래스 분류  
(Multi-class Classification)
3. Softmax 이해
4. pairplot() 함수를 이용한 데이터 분석
5. **Softmax를 이용한 딥러닝 설계 및 구현**

- 딥러닝 설계 및 구현

아이리스 품종 예측을 위한 데이터 → [iris.csv](#)



**TensorBoard를 사용하여 kernel 및 bias 확인하는 방법**

**21\_(17 page) 강의용\_Iris (TensorBoard).ipynb**