

머신러닝과 딥러닝

Report5

소프트웨어학과

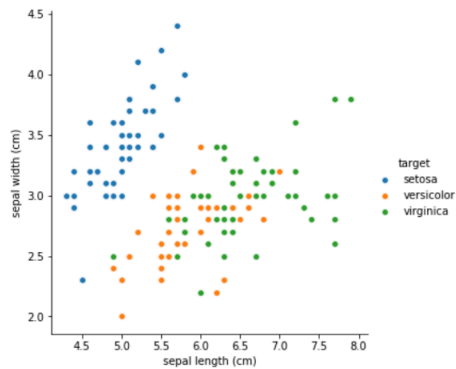
2016312568 정희운

```
In [10]: import seaborn as sn
```

```
In [17]: sn.pairplot(iris_frame,  
                x_vars=["sepal length (cm)"],  
                y_vars=["sepal width (cm)"],  
                hue = "target",  
                size=5)
```

C:\Users\user\Anaconda3\lib\site-packages\seaborn\axisgrid.py:2065: UserWarning: The `size` parameter has been renamed to `height`; please update your code.
warnings.warn(msg, UserWarning)

```
Out[17]: <seaborn.axisgrid.PairGrid at 0x2110277d3c8>
```

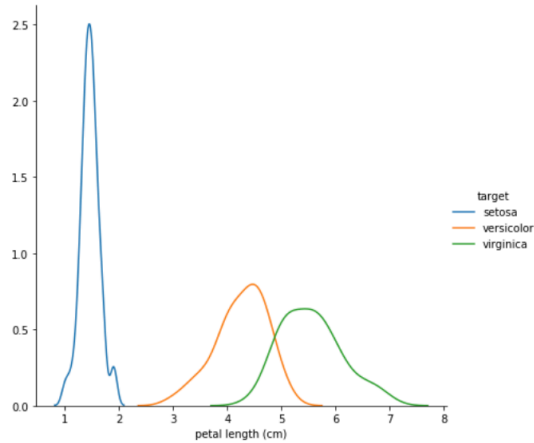


위는 Iris data에서 target의 sepal length, sepal width의 데이터 분포를 나타낸다.

```
In [18]: sn.FacetGrid(iris_frame,
hue="target",
size=6).map(sn.kdeplot, "petal length (cm)").add_legend()

C:\Users\user\Anaconda3\lib\site-packages\seaborn\axisgrid.py:230: UserWarning: The `size` paramter has been
renamed to `height`; please update your code.
warnings.warn(msg, UserWarning)
C:\Users\user\Anaconda3\lib\site-packages\scipy\stats\stats.py:1713: FutureWarning: Using a non-tuple sequen
ce for multidimensional indexing is deprecated; use `arr[tuple(seq)]` instead of `arr[seq]`. In the future t
his will be interpreted as an array index, `arr[np.array(seq)]`, which will result either in an error or a d
ifferent result.
return np.add.reduce(sorted[indexer] * weights, axis=axis) / sumval
```

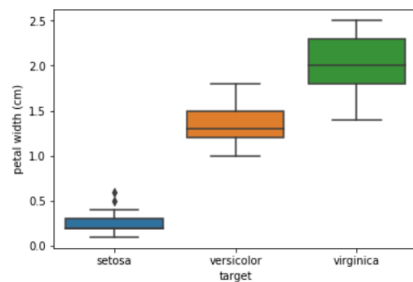
Out[18]: <seaborn.axisgrid.FacetGrid at 0x21103783208>



위는 Iris data에서 target의 petal length의 데이터 분포를 나타낸다.

```
In [19]: sn.boxplot(x="target",y="petal width (cm)",
data=iris_frame)
```

Out[19]: <matplotlib.axes._subplots.AxesSubplot at 0x211038547b8>



위는 Iris data에서 target의 petal width의 데이터 분포를 나타낸다.

```
In [26]: import matplotlib.colors as colors
from sklearn.naive_bayes import GaussianNB
df1 = iris.frame[["sepal length (cm)", "sepal width (cm)", "target" ]]
X = df1.iloc[:,0:2]
Y = df1.iloc[:,2].replace({'setosa':0,'versicolor':1,'virginica':2}).copy()
NB=GaussianNB()
NB.fit(X,Y)
N=100

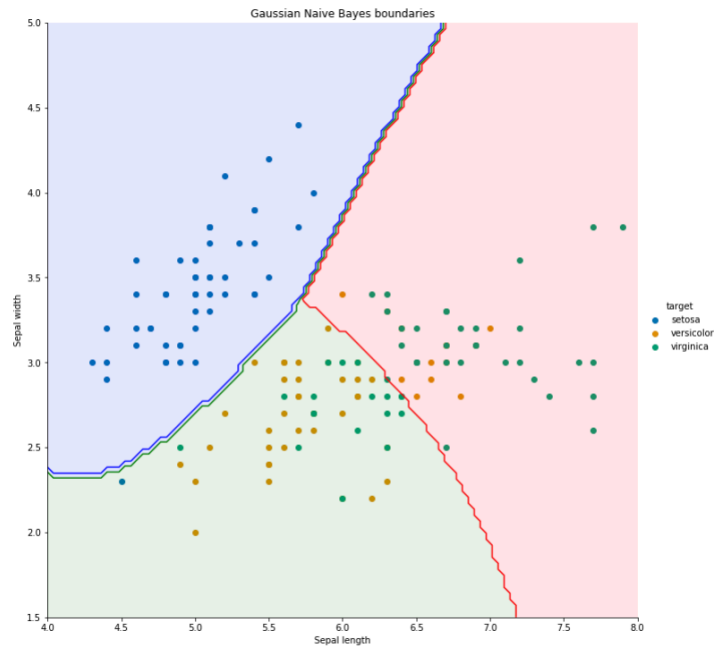
In [31]: X = np.linspace(4,8,N)
Y = np.linspace(1.5,5,N)
X_,Y_ = np.meshgrid(X,Y)
color_list = ['Blues','Greens','Reds']
my_norm = colors.Normalize(vmin=-1,vmax=1)
g = sn.FacetGrid(iris.frame, hue = 'target', size = 10,
                 palette = 'colorblind').map(plt.scatter, "sepal length (cm)", "sepal width (cm)")

my_ax = g.ax
zz = np.array( [NB.predict ([[xx,yy]]) [0] for xx,yy in zip(np.ravel(X_),np.ravel(Y_))])
Z=zz.reshape(X_.shape)
my_ax.contourf(X_,Y_,Z,Z,alpha=.1, colors = ('blue','green','red'))
my_ax.contour(X_,Y_,Z,Z,alpha=.1, colors = ('blue','green','red'))

my_ax.set_xlabel('Sepal length')
my_ax.set_ylabel('Sepal width')
my_ax.set_title('Gaussian Naive Bayes boundaries')

C:\Users\user\Anaconda3\lib\site-packages\seaborn\axisgrid.py:230: UserWarning: The 'size' paramter has been renamed to 'height'; please update your code.
warnings.warn(msg, UserWarning)
```

Out[31]: Text(0.5, 1.0, 'Gaussian Naive Bayes boundaries')



Data model을 GaussianNB를 사용하여 target을 분리하기 위한 선들과 분리된 target을 나타낸 그래프를 출력한 결과이다.