

Exercises 1.

Perform a logistic regression on the preceding data, using three-fold cross-validation and accuracy as your metric.

```
In [46]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import KFold, cross_val_score, train_test_split
from sklearn.metrics import confusion_matrix, classification_report, ConfusionMatrixDisplay
```

```
In [47]: df = pd.read_csv('https://raw.githubusercontent.com/thomasniel/machine-learning/master/data/credit_data.csv')
```

```
In [24]: x = df.values[:, :-1]
y = df.values[:, -1]
```

```
In [25]: kfold = KFold(n_splits=3, random_state=7, shuffle=True)
```

```
In [26]: model = LogisticRegression(penalty=None)
```

```
In [27]: results = cross_val_score(model, x, y, cv=kfold)
```

```
In [28]: results.mean()
```

```
Out[28]: 1.0
```

```
In [29]: results.std()
```

```
Out[29]: 0.0
```

Exercise 2.

Produce a confusion matrix comparing the predictions and actual data.

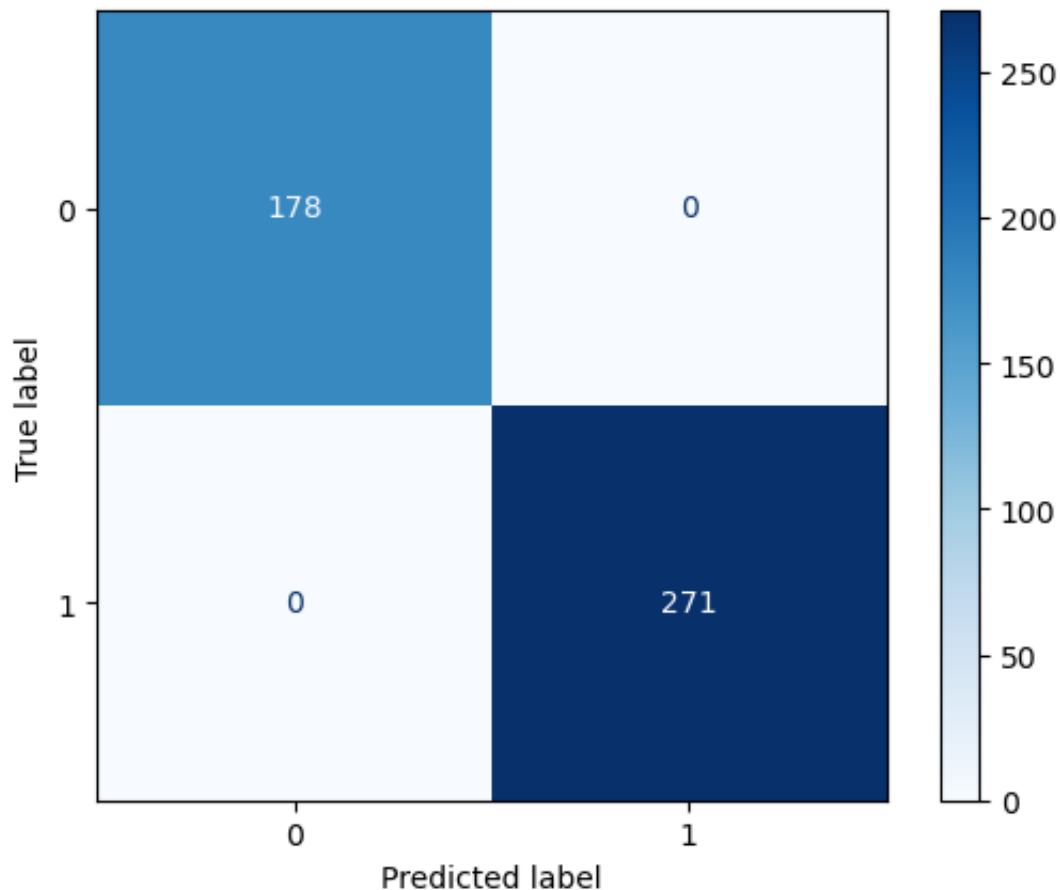
```
In [54]: x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=1/3, shuffle=True)

model.fit(x_train, y_train)
y_pred = model.predict(x_test)

cm = confusion_matrix(y_test, y_pred)
print(classification_report(y_test, y_pred))
```

```
disp = ConfusionMatrixDisplay(confusion_matrix= cm, display_labels=model.classes_)
disp.plot(cmap='Blues')
plt.show();
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	178
1	1.00	1.00	1.00	271
accuracy			1.00	449
macro avg	1.00	1.00	1.00	449
weighted avg	1.00	1.00	1.00	449



Exercise 3

Pick a few different background colors (you can use an RGB tool like this one) and see if the logistic regression sensibly chooses a light (0) or dark (1) font for each one.

```
In [58]: # Test the prediction
n = input('Input a color {red}, {green}, {blue}:')
```

```
(r, g, b) = n.split(',')  
  
x = model.predict(np.array([[int(r), int(g), int(b)]]))  
if x == 0:  
    print('light')  
else:  
    print('dark')
```

dark

Exercie 4

Based on the preceding exercises, do you think logistic regression is effective for predicting a light or dark font for a given background color?

Answer: Yes, the logistic regression is very effective at predicting light or dark fonts for a given background color. Not only is the accuracy extremely high, but the confusion matrix has high numbers in the TP and TN.

In []: