regression hypothesis

May 20, 2021

[2]:
```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
%matplotlib inline
from sklearn.model_selection import train_test_split,cross_val_score
```

[3]:
```
dataset = pd.read_csv('dataset.csv')
X= dataset.drop(columns=['url','status'])
Y= dataset['status']
X.head()
```

[3]:

	length_url	abnormal_subdomain	links_in_tags	submit_email	ratio_intMedia	ratio_extMedia	sfh	iframe	popup_window	safe_anchor	onmouseover	right_clic
0	46	0	73.913043	0	100.000000	0.000000	0	0	0	77.777778	0	0
1	128	0	0.000000	0	0.000000	0.000000	0	0	0	0.000000	0	0
2	52	0	100.000000	0	0.000000	0.000000	0	0	0	0.000000	0	0
3	21	0	100.000000	0	92.307692	7.692308	0	0	0	82.539683	0	0
4	28	0	55.555556	0	50.000000	50.000000	0	0	0	81.081081	0	0

[5]:
```
df1 = pd.DataFrame(X['length_url'])
# df1.iloc[0, 0] = 0
# 1 => legitimate
# -1 => physhing
# 0 => suspecious
```

```
conds = [df1.values < 54 , df1.values > 75]
choices = ['1', '-1']

col = pd.DataFrame(np.select(conds, choices, default='0'),
             index=df1.index,
             columns=df1.columns)

col.head()
```

[5]:
```
  length_url
0          1
1         -1
2          1
3          1
4          1
```

[6]:
```
# append the updated length_url column to remaining dataset
X = X.assign(length_url=col['length_url'])
X.head()
```

[6]:
```
  length_url  abnormal_subdomain  links_in_tags  submit_email  ratio_intMedia  \
0          1                   0      73.913043             0      100.000000
1         -1                   0       0.000000             0        0.000000
2          1                   0     100.000000             0        0.000000
3          1                   0     100.000000             0       92.307692
4          1                   0      55.555556             0       50.000000

   ratio_extMedia  sfh  iframe  popup_window  safe_anchor  onmouseover  \
0        0.000000    0       0             0    77.777778            0
1        0.000000    0       0             0     0.000000            0
2        0.000000    0       0             0     0.000000            0
3        7.692308    0       0             0    82.539683            0
4       50.000000    0       0             0    81.081081            0

   right_clic
0           0
1           0
2           0
3           0
4           0
```

[7]:
```
train_X,test_X,train_Y,test_Y=train_test_split(X,Y,test_size=0.2,random_state=2)
```

[8]:
```
# test using logistic regression
from sklearn.linear_model import LogisticRegression
```

```
from sklearn.metrics import␣
 ↪accuracy_score,confusion_matrix,classification_report
```

[9]:
```
logreg=LogisticRegression()
model_1=logreg.fit(train_X,train_Y)
```

```
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

[10]:
```
logreg_predict= model_1.predict(test_X)
```

[11]:
```
accuracy_score(logreg_predict,test_Y)
```

[11]:
```
0.708125
```

[12]:
```
print(classification_report(logreg_predict,test_Y))
```

```
              precision    recall  f1-score   support

           0       0.69      0.71      0.70       765
           1       0.72      0.71      0.72       835

    accuracy                           0.71      1600
   macro avg       0.71      0.71      0.71      1600
weighted avg       0.71      0.71      0.71      1600
```

0.1 To Plot the Graph Accuracy vs Condition

legitimate value

[41]:
```
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import␣
 ↪accuracy_score,confusion_matrix,classification_report

dataset = pd.read_csv('dataset.csv')
X= dataset.drop(columns=['url','status'])
Y= dataset['status']
X.head()
```

[41]:

	length_url	abnormal_subdomain	links_in_tags	submit_email	\
0	46	0	73.913043	0	
1	128	0	0.000000	0	
2	52	0	100.000000	0	
3	21	0	100.000000	0	
4	28	0	55.555556	0	

	ratio_intMedia	ratio_extMedia	sfh	iframe	popup_window	safe_anchor	\
0	100.000000	0.000000	0	0	0	77.777778	
1	0.000000	0.000000	0	0	0	0.000000	
2	0.000000	0.000000	0	0	0	0.000000	
3	92.307692	7.692308	0	0	0	82.539683	
4	50.000000	50.000000	0	0	0	81.081081	

	onmouseover	right_clic
0	0	0
1	0	0
2	0	0
3	0	0
4	0	0

[97]:
```
lenghth_col = pd.DataFrame(X['length_url'])

graph_data = pd.DataFrame()

graph_data["url_length"] = range(75)

# graph_data["accuracy"] = [76

accuracy_arr = ['none']*75

for num1 in range(75):
    conds = [lenghth_col.values < num1 , lenghth_col.values > 75]
    choices = ['1', '-1']

    col = pd.DataFrame(np.select(conds, choices, default='0'),
                index=df1.index,
                columns=df1.columns)

    # append the updated length_url column to remaining dataset
    X_new = X.assign(length_url=col['length_url'])

    train_X,test_X,train_Y,test_Y=train_test_split(X_new,Y,test_size=0.
↪2,random_state=2)

    # test using logistic regression
    logreg=LogisticRegression()
```

```
    model_1=logreg.fit(train_X,train_Y)

    logreg_predict= model_1.predict(test_X)
    accuracy_arr[num1] = accuracy_score(logreg_predict,test_Y)*100



graph_data["accuracy"] = accuracy_arr
```

```
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
```

```
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
```

```
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
```

```
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
```

```
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

```
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.
```

```
Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

[98]:
```
graph_data.head()
```

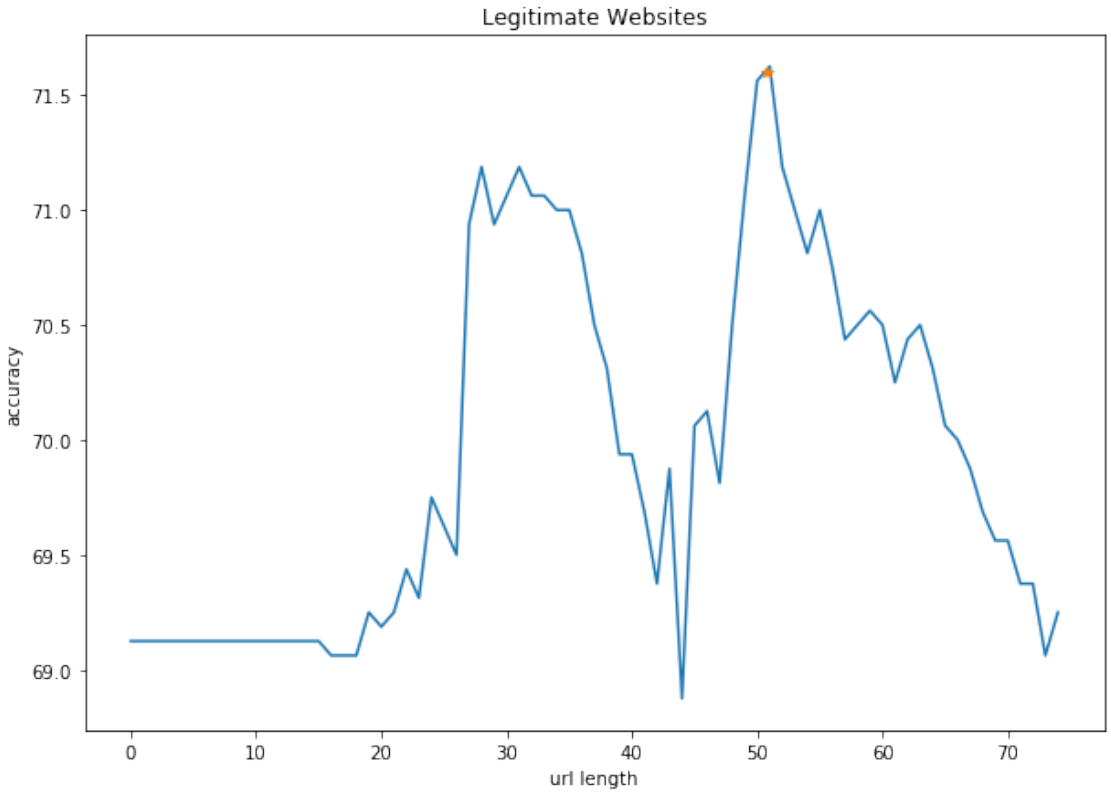
[98]:

	url_length	accuracy
0	0	69.125
1	1	69.125
2	2	69.125
3	3	69.125
4	4	69.125

[104]:
```
import matplotlib.pyplot as plt
%matplotlib inline

plt.figure(figsize=(10, 7))
plt.title('Legitimate Websites')
plt.xlabel('url length')
plt.ylabel('accuracy')

plt.plot(graph_data['url_length'].values, graph_data['accuracy'].values )
plt.plot(50.8,71.6, '*')
plt.show()
```



According to the above graph we can assume that url lengths less than 51 most likely to be legitimate

0.1.1 phishing websites

[106]:
```
lenghth_col = pd.DataFrame(X['length_url'])

graph_data = pd.DataFrame()

graph_data["url_length"] = range(52,100)

graph_data.head()


accuracy_arr = ['none']*48
# index =0

for num2 in range(52,100):
    conds = [lenghth_col.values < 51 , lenghth_col.values > num2]
    choices = ['1', '-1']

    col = pd.DataFrame(np.select(conds, choices, default='0'),
                index=df1.index,
```

```
                columns=df1.columns)

    # append the updated length_url column to remaining dataset
    X_new = X.assign(length_url=col['length_url'])

    train_X,test_X,train_Y,test_Y=train_test_split(X_new,Y,test_size=0.
↪2,random_state=2)

    # test using logistic regression
    logreg=LogisticRegression()
    model_1=logreg.fit(train_X,train_Y)

    logreg_predict= model_1.predict(test_X)
    accuracy_arr[num2-52] = accuracy_score(logreg_predict,test_Y)*100



graph_data["accuracy"] = accuracy_arr
```

```
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
```

```
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
```

```
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
```

```
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
```

```
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

```
c:\users\lakru\appdata\local\programs\python\python37\lib\site-
packages\sklearn\linear_model\_logistic.py:765: ConvergenceWarning: lbfgs failed
to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-
regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

[107]:
```
graph_data.head()
```

[107]:

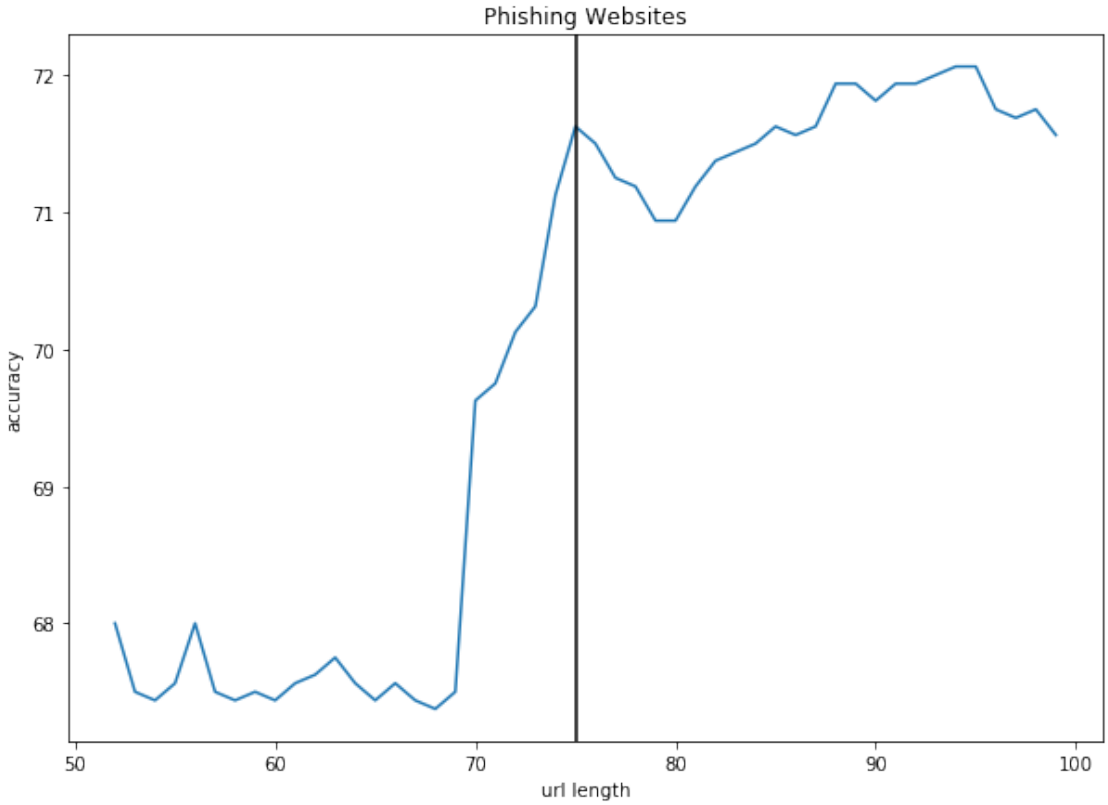
	url_length	accuracy
0	52	68.0000
1	53	67.5000
2	54	67.4375
3	55	67.5625
4	56	68.0000

[108]:
```
import matplotlib.pyplot as plt
%matplotlib inline

plt.figure(figsize=(10, 7))


plt.plot(graph_data['url_length'].values, graph_data['accuracy'].values )
plt.title('Phishing Websites')
plt.xlabel('url length')
plt.ylabel('accuracy')


plt.axvline(x=75, color='black')
plt.show()
```



According to the above graph we can assume that url lengths higher than 75 are most likely to be phishing

[ ]: