Name : Kshitij V Darwhekar

Roll No: TETB19

Sub: Soft Computitng

Batch: B2

3. RANDOM FOREST

```
import pandas as pd
from sklearn.datasets import load_digits
digits = load_digits()
```

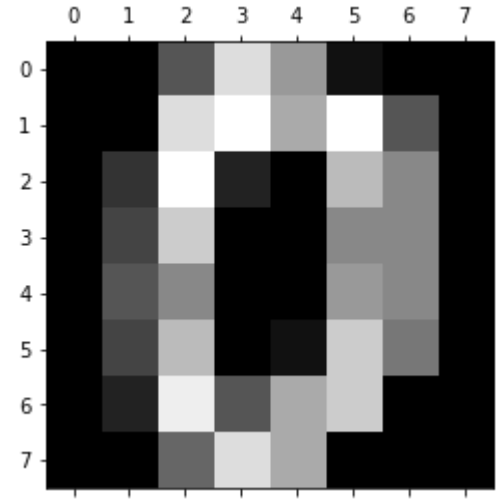
```
dir(digits)
```

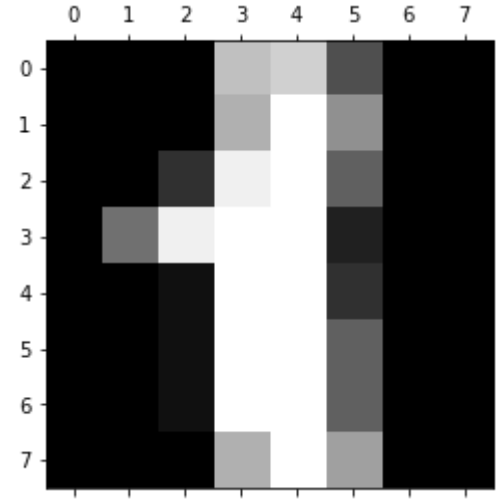
```
['DESCR', 'data', 'feature_names', 'frame', 'images', 'target', 'target_names']
```

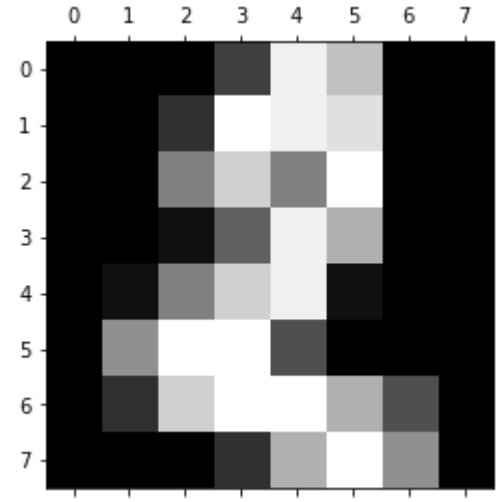
```
%matplotlib inline
import matplotlib.pyplot as plt
```

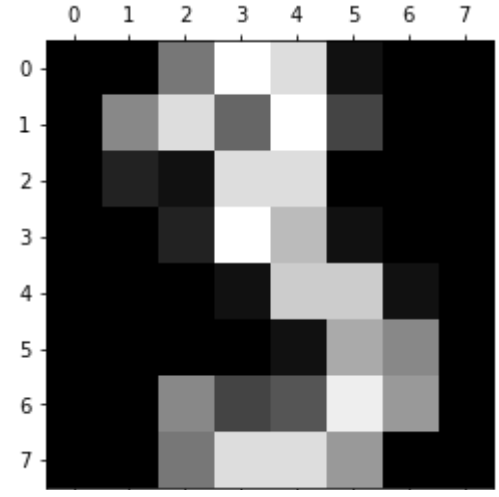
```
plt.gray()
for i in range(10):
  plt.matshow(digits.images[i])
```

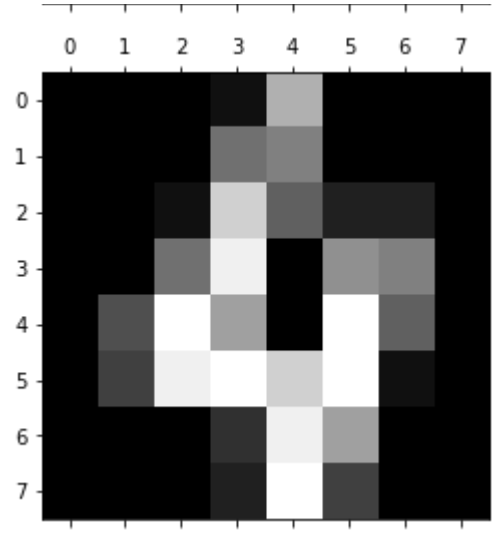
<Figure size 432x288 with 0 Axes>

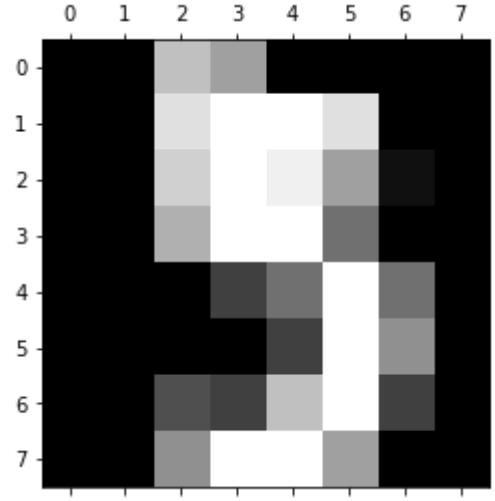


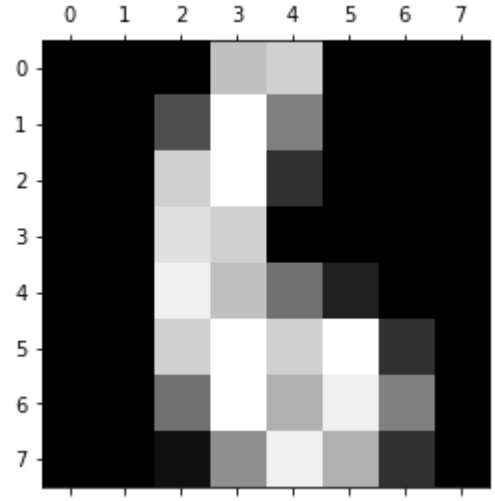


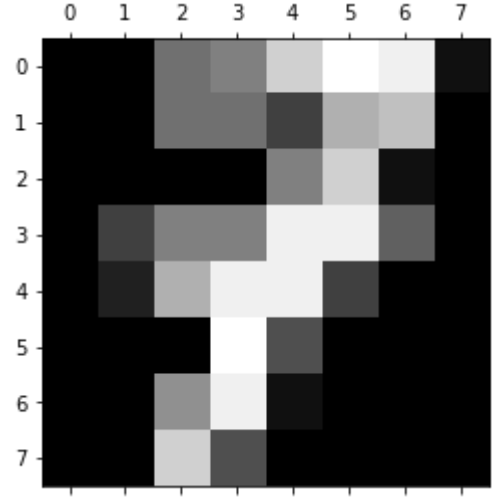


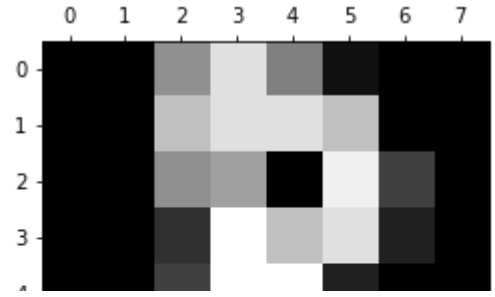



0 1 2 3 4 5 6 7
0
1
2
3
4
5
6
7


0 1 2 3 4 5 6 7
0
1
2
3
4
5
6
7


0 1 2 3 4 5 6 7
0
1
2
3
4
5
6
7


0 1 2 3 4 5 6 7
0
1
2
3
4
5
6
7




```
df = pd.DataFrame(digits.data)
df.head()
```

	0	1	2	3	4	5	6	7	8	9	...	54	55	56	57	58	59	
0	0.0	0.0	5.0	13.0	9.0	1.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	6.0	13.0	1
1	0.0	0.0	0.0	12.0	13.0	5.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	11.0	1
2	0.0	0.0	0.0	4.0	15.0	12.0	0.0	0.0	0.0	0.0	...	5.0	0.0	0.0	0.0	0.0	3.0	1
3	0.0	0.0	7.0	15.0	13.0	1.0	0.0	0.0	0.0	8.0	...	9.0	0.0	0.0	0.0	7.0	13.0	1
4	0.0	0.0	0.0	1.0	11.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	2.0	1

5 rows × 64 columns

```
df['target'] = digits.target
```

```
df[0:12]
```

	0	1	2	3	4	5	6	7	8	9	...	55	56	57	58	59	
0	0.0	0.0	5.0	13.0	9.0	1.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	6.0	13.0	1
1	0.0	0.0	0.0	12.0	13.0	5.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	11.0	1
2	0.0	0.0	0.0	4.0	15.0	12.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	3.0	1
3	0.0	0.0	7.0	15.0	13.0	1.0	0.0	0.0	0.0	8.0	...	0.0	0.0	0.0	7.0	13.0	1
4	0.0	0.0	0.0	1.0	11.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	2.0	1
5	0.0	0.0	12.0	10.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	9.0	16.0	1
6	0.0	0.0	0.0	12.0	13.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	1.0	9.0	1
7	0.0	0.0	7.0	8.0	13.0	16.0	15.0	1.0	0.0	0.0	...	0.0	0.0	0.0	13.0	5.0	
8	0.0	0.0	9.0	14.0	8.0	1.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	11.0	16.0	1
9	0.0	0.0	11.0	12.0	0.0	0.0	0.0	0.0	0.0	2.0	...	0.0	0.0	0.0	9.0	12.0	1
10	0.0	0.0	1.0	9.0	15.0	11.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	1.0	10.0	1
11	0.0	0.0	0.0	0.0	14.0	13.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	1.0	1

12 rows × 65 columns

```
## Train the model and prediction
```

```
X = df.drop('target',axis = 'columns')
y = df.target
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.1)
```

```
from sklearn.ensemble import RandomForestClassifier
model = RandomForestClassifier(n_estimators=30)
model.fit(X_train, y_train)
```

```
     RandomForestClassifier(n_estimators=30)
```

```
model.score(X_test, y_test)
```

```
     0.95
```

```
y_predicted = model.predict(X_test)
```

```
from sklearn.datasets import make_classification
```

```
## Confusion Matrix
```

```
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_predicted)
cm
```

```
     array([[18,  0,  0,  0,  0,  0,  0,  0,  0,  0],
            [ 0, 17,  0,  0,  0,  0,  0,  0,  0,  0],
            [ 0,  0, 15,  0,  0,  0,  0,  0,  0,  0],
            [ 0,  0,  0, 25,  0,  0,  0,  0,  0,  0],
            [ 0,  0,  0,  0, 15,  0,  0,  0,  0,  0],
            [ 0,  0,  0,  0,  1, 18,  1,  0,  0,  1],
            [ 0,  0,  0,  0,  0,  0, 18,  0,  0,  0],
            [ 0,  0,  0,  0,  0,  0,  0, 15,  0,  0],
            [ 0,  0,  1,  2,  0,  0,  0,  0, 14,  0],
            [ 0,  1,  0,  0,  0,  1,  0,  0,  1, 16]])
```

```
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sn
plt.figure(figsize=(10,7))
sn.heatmap(cm, annot=True)
plt.xlabel('Predicted')
plt.ylabel('Truth')
```