DIGIT RECOGNIZER USING DECISION TREE

BY

R THARUN KUMAR

V SAI TEJA

T PRUDHVI KRISHNA

B VAMSHIKRISHNA

In [1]:

```
#importing all libraries
```

In [2]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as pt
from sklearn.tree import DecisionTreeClassifier
```

In [3]:

```
# loading data file using pandas
```

In [4]:

```
data = pd.read_csv('C://Users//bandari vamshi//Desktop//ml pro//train.csv').as_matrix()
```

```
O:\Anaconda_Files\lib\site-packages\ipykernel_launcher.py:1: FutureWarnin
g: Method .as_matrix will be removed in a future version. Use .values inst
ead.
  """Entry point for launching an IPython kernel.
```

In [5]:

```
# Taking decision tree classification model from scikit-learn
```

In [6]:

```
model = DecisionTreeClassifier()
```

In [7]:

```
# Dividing data into training and testing parts

#first coloumn consists of label
```

In [8]:

```
train_data = data[0:21000,1:]
train_label = data[0:21000,:1]
```

In [9]:

```
test_data = data[21000:,1:]
test_label = data[21000:, :1]
```

In [10]:

```
#train model by passing training data and label data
```

In [11]:

```
model.fit(train_data, train_label)
```

Out[11]:

```
DecisionTreeClassifier(class_weight=None, criterion='gini', max_depth=Non
e,
                       max_features=None, max_leaf_nodes=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, presort=False,
                       random_state=None, splitter='best')
```

In [12]:

```
#predicting the model and storing output values in list a
```

In [13]:

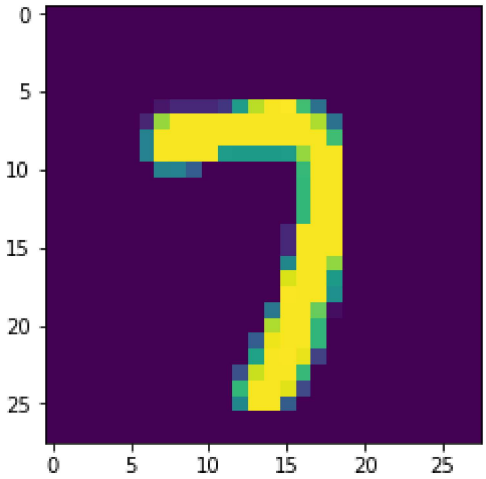
```
a = model.predict(test_data)
```

In [14]:

```
# plotting image of digit
```

In [15]:

```
res = test_data[305]
res.shape = (28,28)
pt.imshow(res)
pt.show()
```



In [16]:

```
#printing the value of above image
```

In [17]:

```
a[305]
```

Out[17]:

7

In [ ]:

```

```