In [ ]:
```
from sklearn.datasets import load_iris
from sklearn import tree
import pandas as pd
import matplotlib.pyplot as plt
```

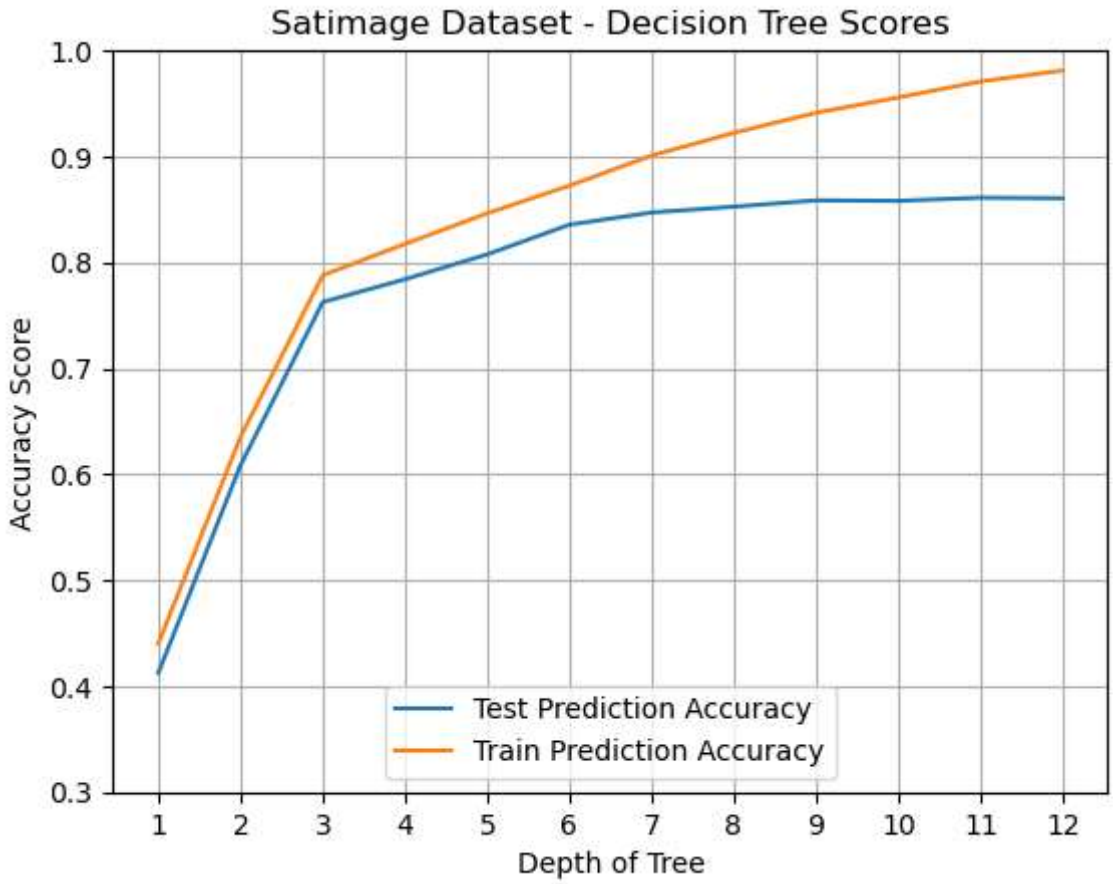
In [ ]:
```
X = pd.read_table('D:\PythonS\DataSets\FSU Applied Machine Learning DS\satimage\X.dat'
Y = pd.read_table('D:\PythonS\DataSets\FSU Applied Machine Learning DS\satimage\Y.dat'
```

In [ ]:
```
trees = []
for i in range(1,13):
    trees.append(tree.DecisionTreeClassifier(max_depth=i, random_state= i).fit(X,Y))
```

In [ ]:
```
Xtest = pd.read_table('D:\PythonS\DataSets\FSU Applied Machine Learning DS\satimage\Xt
Ytest = pd.read_table('D:\PythonS\DataSets\FSU Applied Machine Learning DS\satimage\Yt
```

In [ ]:
```
train_scores = []
test_scores = []
for t in trees:
    train_scores.append(t.score(X,Y))
    test_scores.append(t.score(Xtest,Ytest))
```

In [ ]:
```
plt.title('Satimage Dataset - Decision Tree Scores')
plt.plot(range(1,13), test_scores, label = 'Test Prediction Accuracy')
plt.plot(range(1,13), train_scores, label = 'Train Prediction Accuracy')
plt.ylim(0.3,1)
plt.xlabel('Depth of Tree')
plt.ylabel('Accuracy Score')
plt.legend(loc = 'lower center')
plt.xticks(range(1,13))
plt.yticks()
plt.grid()
plt.show()
```



In [ ]:
```
print(1 - test_scores[-1])
```

```
0.13949999999999996
```

In [ ]: