Allen Ben Philipose - 18BIS0043

Lab FAT, L53+L54

ECE3502 - IoT Domain Analyst

Submitted to: Prof. Pradheep T

Aim

In the given dataset, use random forest and predict the number of cases for

- 01.03.2021,Belgium (Reg.No: 18BIS0043)

Algorithm Explanation

Random forests, or random decision forests, are an ensemble learning technique for classification, regression, and other tasks. It works by training a large number of decision trees and then outputting the class that is the mode of the classes (classification) or the mean/average prediction (regression) of the individual trees. Random decision forests compensate for decision trees' inclination for excessive fit to their training set. Although random forests surpass decision trees in practice, their accuracy is lower than that of gradient boosted trees. However, the features of the data might have an effect on their performance.

Each tree in the random forest generates a class prediction, and the class with the most votes becomes the prediction of our model. A large number of substantially uncorrelated models (trees) acting in conjunction outperforms any of the component models individually. A random forest is a meta estimator that fits a number of decision tree classifiers to different subsamples of the dataset and then utilises averaging to increase predicted accuracy and avoid overfitting.

"Random Forest is a classifier that combines many decision trees on diverse subsets of a given dataset and uses the average to increase the dataset's predicted accuracy." Rather of depending on a single decision tree, the random forest collects the forecasts from each tree and forecasts the final output based on the majority vote of forecasts.

Info from documentations

Code and Result

Import Libraries

In [53]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

In [54]:
```
from sklearn import preprocessing
from sklearn.metrics import accuracy_score, log_loss, confusion_matrix, f1_score
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.model_selection import StratifiedKFold, GridSearchCV, KFold
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import classification_report
```

In [55]:
```
from sklearn.ensemble import RandomForestClassifier
from sklearn.ensemble import RandomForestRegressor
from sklearn.tree import DecisionTreeClassifier
from sklearn.naive_bayes import GaussianNB
from sklearn.neighbors import KNeighborsClassifier
from sklearn.linear_model import LogisticRegression
```

In [56]:
```
import warnings
warnings.filterwarnings('ignore')
```

In [57]:
```
from sklearn.metrics import accuracy_score
from sklearn.metrics import mean_squared_error
from scipy.stats import spearmanr
```

Dataset

In [202]:
```
allen = pd.read_csv("Dataset.csv")
```

In [203]:
```
allen.head()
```

Out[203]:

	dateRep	day	month	year	cases	deaths	countriesAndTerritories	geoId	countryterritoryCode	popData2019	continentExp	Cumulati
0	14-12-2020	14	12	2020	746	6	Afghanistan	AF	AFG	38041757.0	Asia	
1	13-12-2020	13	12	2020	298	9	Afghanistan	AF	AFG	38041757.0	Asia	
2	12-12-2020	12	12	2020	113	11	Afghanistan	AF	AFG	38041757.0	Asia	
3	11-12-2020	11	12	2020	63	10	Afghanistan	AF	AFG	38041757.0	Asia	
4	10-12-2020	10	12	2020	202	16	Afghanistan	AF	AFG	38041757.0	Asia	

Pre-processing

In [204]:
```
allen = allen.drop_duplicates()
```

In [205]:
```
allen = allen.dropna()
```

In [206]:
```
allen.isnull().sum()
```

Out[206]:
```
dateRep                                            0
day                                                0
month                                              0
year                                               0
cases                                              0
deaths                                             0
countriesAndTerritories                            0
geoId                                              0
countryterritoryCode                               0
popData2019                                        0
continentExp                                       0
Cumulative_number_for_14_days_of_COVID-19_cases_per_100000    0
dtype: int64
```

In [207]:
```
allen = allen[allen['continentExp'] == 'Europe']
allen
```

Out[207]:

	dateRep	day	month	year	cases	deaths	countriesAndTerritories	geoId	countryterritoryCode	popData2019	continentExp	Cum
340	14-12-2020	14	12	2020	788	14	Albania	AL	ALB	2862427.0	Europe	
341	13-12-2020	13	12	2020	879	12	Albania	AL	ALB	2862427.0	Europe	
342	12-12-2020	12	12	2020	802	12	Albania	AL	ALB	2862427.0	Europe	
343	11-12-2020	11	12	2020	873	14	Albania	AL	ALB	2862427.0	Europe	
344	10-12-2020	10	12	2020	752	15	Albania	AL	ALB	2862427.0	Europe	
...	...	...	...	...	...	...	...	...	...	...	...	
58710	17-01-2020	17	1	2020	0	0	United_Kingdom	UK	GBR	66647112.0	Europe	
58711	16-01-2020	16	1	2020	0	0	United_Kingdom	UK	GBR	66647112.0	Europe	
58712	15-01-2020	15	1	2020	0	0	United_Kingdom	UK	GBR	66647112.0	Europe	
58713	14-01-2020	14	1	2020	0	0	United_Kingdom	UK	GBR	66647112.0	Europe	
58714	13-01-2020	13	1	2020	0	0	United_Kingdom	UK	GBR	66647112.0	Europe	

16657 rows × 12 columns

In [208]:
```
allen.head()
```

Out[208]:

	dateRep	day	month	year	cases	deaths	countriesAndTerritories	geoId	countryterritoryCode	popData2019	continentExp	Cumul
340	14-12-2020	14	12	2020	788	14	Albania	AL	ALB	2862427.0	Europe	
341	13-12-2020	13	12	2020	879	12	Albania	AL	ALB	2862427.0	Europe	
342	12-12-2020	12	12	2020	802	12	Albania	AL	ALB	2862427.0	Europe	
343	11-12-2020	11	12	2020	873	14	Albania	AL	ALB	2862427.0	Europe	
344	10-12-2020	10	12	2020	752	15	Albania	AL	ALB	2862427.0	Europe	

In [215]:
```
allen.drop(allen.columns[4], axis=1, inplace=True)
allen
```

Out[215]:

	day	month	year	cases	countriesAndTerritories
340	14	12	2020	788	Albania
341	13	12	2020	879	Albania
342	12	12	2020	802	Albania
343	11	12	2020	873	Albania
344	10	12	2020	752	Albania
...	...	...	...	...	...
58710	17	1	2020	0	United_Kingdom
58711	16	1	2020	0	United_Kingdom
58712	15	1	2020	0	United_Kingdom
58713	14	1	2020	0	United_Kingdom
58714	13	1	2020	0	United_Kingdom

16657 rows × 5 columns

In [250]:
```
# Backup

allen = alvin
```

In [251]:
```
allen
```

Out[251]:

	day	month	year	cases	countriesAndTerritories
340	14	12	2020	788	Albania
341	13	12	2020	879	Albania
342	12	12	2020	802	Albania
343	11	12	2020	873	Albania
344	10	12	2020	752	Albania
...	...	...	...	...	...
58710	17	1	2020	0	United_Kingdom
58711	16	1	2020	0	United_Kingdom
58712	15	1	2020	0	United_Kingdom
58713	14	1	2020	0	United_Kingdom
58714	13	1	2020	0	United_Kingdom

16657 rows × 5 columns

Train-test-split

In [252]:
```
allen = pd.get_dummies(allen)
```

In [253]:
```
allen
```

Out[253]:

	day	month	year	cases	countriesAndTerritories_Albania	countriesAndTerritories_Andorra	countriesAndTerritories_Armenia
340	14	12	2020	788	1	0	0
341	13	12	2020	879	1	0	0
342	12	12	2020	802	1	0	0
343	11	12	2020	873	1	0	0
344	10	12	2020	752	1	0	0
...	...	...	...	...	...	...	...
58710	17	1	2020	0	0	0	0
58711	16	1	2020	0	0	0	0
58712	15	1	2020	0	0	0	0
58713	14	1	2020	0	0	0	0
58714	13	1	2020	0	0	0	0

16657 rows × 59 columns

In [ ]:
```
xtrain = allen.drop(columns=["cases"])
ytrain = allen["cases"]
```

In [255]:
```
print(xtrain)
print(len(xtrain))
```

```
       day  month  year  countriesAndTerritories_Albania  \
340     14     12  2020                                1
341     13     12  2020                                1
342     12     12  2020                                1
343     11     12  2020                                1
344     10     12  2020                                1
...    ...    ...   ...                              ...
58710   17      1  2020                                0
58711   16      1  2020                                0
58712   15      1  2020                                0
58713   14      1  2020                                0
58714   13      1  2020                                0

       countriesAndTerritories_Andorra  countriesAndTerritories_Armenia  \
340                                  0                                0
341                                  0                                0
342                                  0                                0
343                                  0                                0
344                                  0                                0
...                                ...                              ...
58710                                0                                0
58711                                0                                0
58712                                0                                0
58713                                0                                0
58714                                0                                0

       countriesAndTerritories_Austria  countriesAndTerritories_Azerbaijan  \
340                                  0                                   0
341                                  0                                   0
342                                  0                                   0
343                                  0                                   0
344                                  0                                   0
...                                ...                                 ...
58710                                0                                   0
58711                                0                                   0
58712                                0                                   0
58713                                0                                   0
58714                                0                                   0

       countriesAndTerritories_Belarus  countriesAndTerritories_Belgium  ...  \
340                                  0                                0  ...
341                                  0                                0  ...
342                                  0                                0  ...
343                                  0                                0  ...
344                                  0                                0  ...
...                                ...                              ...  ...
58710                                0                                0  ...
58711                                0                                0  ...
58712                                0                                0  ...
58713                                0                                0  ...
58714                                0                                0  ...

       countriesAndTerritories_San_Marino  countriesAndTerritories_Serbia  \
340                                     0                               0
341                                     0                               0
342                                     0                               0
343                                     0                               0
344                                     0                               0
...                                   ...                             ...
58710                                   0                               0
58711                                   0                               0
58712                                   0                               0
58713                                   0                               0
58714                                   0                               0

       countriesAndTerritories_Slovakia  countriesAndTerritories_Slovenia  \
340                                   0                                 0
341                                   0                                 0
342                                   0                                 0
343                                   0                                 0
344                                   0                                 0
...                                 ...                               ...
58710                                 0                                 0
58711                                 0                                 0
58712                                 0                                 0
58713                                 0                                 0
58714                                 0                                 0

       countriesAndTerritories_Spain  countriesAndTerritories_Sweden  \
340                                0                               0
341                                0                               0
342                                0                               0
343                                0                               0
344                                0                               0
...                              ...                             ...
58710                              0                               0
58711                              0                               0
58712                              0                               0
58713                              0                               0
58714                              0                               0

       countriesAndTerritories_Switzerland  countriesAndTerritories_Turkey  \
340                                      0                               0
341                                      0                               0
342                                      0                               0
343                                      0                               0
344                                      0                               0
...                                    ...                             ...
58710                                    0                               0
58711                                    0                               0
58712                                    0                               0
58713                                    0                               0
58714                                    0                               0

       countriesAndTerritories_Ukraine  countriesAndTerritories_United_Kingdom
340                                  0                                       0
341                                  0                                       0
342                                  0                                       0
343                                  0                                       0
344                                  0                                       0
...                                ...                                     ...
58710                                0                                       1
58711                                0                                       1
58712                                0                                       1
58713                                0                                       1
58714                                0                                       1

[16657 rows x 58 columns]
16657
```

In [256]:
```
data = pd.DataFrame([X_train.iloc[0]])
for col in data.columns:
    data[col].values[:] = 0
data['countriesAndTerritories_Belgium'] = 1
data['day'] = 1
data['month'] = 3
data['year'] = 2021
x_test = data
print(x_test)
```

```
     day  month  year  countriesAndTerritories_Albania  \
340    1      3  2021                                0

     countriesAndTerritories_Andorra  countriesAndTerritories_Armenia  \
340                                0                                0

     countriesAndTerritories_Austria  countriesAndTerritories_Azerbaijan  \
340                                0                                   0

     countriesAndTerritories_Belarus  countriesAndTerritories_Belgium  ...  \
340                                0                                1  ...

     countriesAndTerritories_San_Marino  countriesAndTerritories_Serbia  \
340                                   0                               0

     countriesAndTerritories_Slovakia  countriesAndTerritories_Slovenia  \
340                                 0                                 0

     countriesAndTerritories_Spain  countriesAndTerritories_Sweden  \
340                              0                               0

     countriesAndTerritories_Switzerland  countriesAndTerritories_Turkey  \
340                                    0                               0

     countriesAndTerritories_Ukraine  countriesAndTerritories_United_Kingdom
340                                0                                       0

[1 rows x 58 columns]
```

In [260]:
```
model0 = RandomForestRegressor(n_estimators=20, random_state=0)
model0.fit(X_traina, Y_traina)
y_pred = model0.predict(x_test)
print(y_pred)
```

```
[16.3]
```

Model and Prediction (18BIS0043)

In [257]:
```
data = pd.DataFrame([xtrain.iloc[0]])
for col in data.columns:
    data[col].values[:] = 0
```

In [258]:
```
data['countriesAndTerritories_Belgium'] = 1
data['day'] = 1
data['month'] = 3
data['year'] = 2021
xtest = data
```

Trial I

In [245]:
```
model0 = RandomForestRegressor()
model0.fit(xtrain,ytrain)
p0 = model0.predict(xtest)
```

In [246]:
```
allen2 = pd.DataFrame(columns=['day','month','year','accuracy'])
```

In [247]:
```
values = {"day": 1,
          "month": 3,
          "year": 2021,
          "accuracy": "14.5",
}
allen2 = allen2.append(values, ignore_index=True)
```

In [248]:
```
allen2.head()
```

Out[248]:

	day	month	year	accuracy
0	1	3	2021	

Inferences

We have successfully predicted the number of corona cases occuring on the given date in the given country. The accuracy was drastically less because of two main reasons

- The date being further away from the expected date
- The variance available in the dataset

Besides these, regression algorithms could not relate with the concept of date being exhibited in the integer format hence the inaccuracies continue to increase drastically. Hence there's a huge difference between the actual numbers and the predicted numbers.

This along with the data being perfectly lying on the verge of second wave varied the real value to be over 700 when the predicted value was under 20.

Viva

Explain the differences between data lake and data warehousing? Which is most suitable for business intelligence applications?

Data warehouse is most suitable for business intelligence application due to the variety of applications that can be made from them due to the large amount of raw data available.

In terms of data storage, the data lake stores all forms of data in their raw form, including structured and unstructured data. It contains data that may be valuable in the present but is also likely to be valuable in the future. Data warehouse only contains high-quality data that has already been pre-processed and is ready for the team to use. In terms of users, Data lakes are used by data scientists to discover patterns and relevant information that can aid companies, while data warehouses are used by business analysts to develop visualizations and reports. In terms of purpose, the purpose of Data Lake is not fixed. Sometimes organizations have a future use-case in mind. Its general uses include data discovery, user profiling, and machine learning. The data warehouse contains data that has been pre-designed for a certain use case. Business Intelligence, Visualizations, and Batch Reporting are just a few of the applications. In terms of cost, data lakes are comparatively cheaper as data is not necessarily stored in a structured manner. Data warehouses on the other hand are more expensive as well as time consuming.

Trial II

Random Forest

In [97]:
```
model1 = RandomForestRegressor()
model1.fit(xtrain,ytrain)
p1 = model1.predict(xtest)
```

In [98]:
```
print(confusion_matrix(ytest,p1))
```

```
[[103   1   2 ...   0   0   0]
 [  8   3   0 ...   0   0   0]
 [  3   5   0 ...   0   0   0]
 ...
 [  0   0   0 ...   0   0   0]
 [  0   0   0 ...   0   0   0]
 [  0   0   0 ...   0   1   0]]
```

In [100]:
```
print(classification_report(ytest.values.ravel(),p1,target_names=list(str(int) for int in range(0,722
))))
```

```
        0.00         0
         355       0.00      0.00      0.00         0
         356       0.00      0.00      0.00         1
         357       0.00      0.00      0.00         0
         358       0.00      0.00      0.00         0
         359       0.00      0.00      0.00         0
         360       0.00      0.00      0.00         2
         361       0.00      0.00      0.00         1
         362       0.00      0.00      0.00         1
         363       0.00      0.00      0.00         0
         364       0.00      0.00      0.00         1
         365       0.00      0.00      0.00         0
         366       0.00      0.00      0.00         0
         367       0.00      0.00      0.00         1
         368       0.00      0.00      0.00         0
         369       0.00      0.00      0.00         1
         370       0.00      0.00      0.00         1
         371       0.00      0.00      0.00         1
         372       0.00      0.00      0.00         1
         373       0.00      0.00      0.00         0
         374       0.00      0.00      0.00         0
         375       0.00      0.00      0.00         1
         376       0.00      0.00      0.00         0
         377       0.00      0.00      0.00         1
         378       0.00      0.00      0.00         1
         379       0.00      0.00      0.00         1
         380       0.00      0.00      0.00         1
         381       0.00      0.00      0.00         1
         382       0.00      0.00      0.00         0
         383       0.00      0.00      0.00         0
         384       0.00      0.00      0.00         0
         385       0.00      0.00      0.00         0
         386       0.00      0.00      0.00         1
         387       0.00      0.00      0.00         0
         388       0.00      0.00      0.00         0
         389       0.00      0.00      0.00         1
         390       0.00      0.00      0.00         1
         391       0.00      0.00      0.00         0
         392       0.00      0.00      0.00         1
         393       0.00      0.00      0.00         1
         394       0.00      0.00      0.00         2
         395       0.00      0.00      0.00         1
         396       0.00      0.00      0.00         1
         397       0.00      0.00      0.00         1
         398       0.00      0.00      0.00         1
         399       0.00      0.00      0.00         1
         400       0.00      0.00      0.00         0
         401       0.00      0.00      0.00         0
         402       0.00      0.00      0.00         0
         403       0.00      0.00      0.00         2
         404       0.00      0.00      0.00         0
         405       0.00      0.00      0.00         0
         406       0.00      0.00      0.00         1
         407       0.00      0.00      0.00         1
         408       0.00      0.00      0.00         1
         409       0.00      0.00      0.00         0
         410       0.00      0.00      0.00         0
         411       0.00      0.00      0.00         1
         412       0.00      0.00      0.00         1
         413       0.00      0.00      0.00         0
         414       0.00      0.00      0.00         0
         415       0.00      0.00      0.00         0
         416       0.00      0.00      0.00         0
         417       0.00      0.00      0.00         0
         418       0.00      0.00      0.00         1
         419       0.00      0.00      0.00         0
         420       0.00      0.00      0.00         0
         421       0.00      0.00      0.00         1
         422       0.00      0.00      0.00         1
         423       0.00      0.00      0.00         0
         424       0.00      0.00      0.00         0
         425       0.00      0.00      0.00         1
         426       0.00      0.00      0.00         0
         427       0.00      0.00      0.00         0
         428       0.00      0.00      0.00         0
         429       0.00      0.00      0.00         1
         430       0.00      0.00      0.00         0
         431       0.00      0.00      0.00         1
         432       0.00      0.00      0.00         0
         433       0.00      0.00      0.00         0
         434       0.00      0.00      0.00         0
         435       0.00      0.00      0.00         0
         436       0.00      0.00      0.00         1
         437       0.00      0.00      0.00         1
         438       0.00      0.00      0.00         1
         439       0.00      0.00      0.00         1
         440       0.00      0.00      0.00         1
         441       0.00      0.00      0.00         1
         442       0.00      0.00      0.00         1
         443       0.00      0.00      0.00         0
         444       0.00      0.00      0.00         1
         445       0.00      0.00      0.00         0
         446       0.00      0.00      0.00         1
         447       0.00      0.00      0.00         1
         448       0.00      0.00      0.00         1
         449       0.00      0.00      0.00         0
         450       0.00      0.00      0.00         0
         451       0.00      0.00      0.00         0
         452       0.00      0.00      0.00         0
         453       0.00      0.00      0.00         0
         454       0.00      0.00      0.00         2
         455       0.00      0.00      0.00         1
         456       0.00      0.00      0.00         0
         457       0.00      0.00      0.00         2
         458       0.00      0.00      0.00         1
         459       0.00      0.00      0.00         0
         460       0.00      0.00      0.00         0
         461       0.00      0.00      0.00         0
         462       0.00      0.00      0.00         1
         463       0.00      0.00      0.00         0
         464       0.00      0.00      0.00         0
         465       0.00      0.00      0.00         0
         466       0.00      0.00      0.00         0
         467       0.00      0.00      0.00         1
         468       0.00      0.00      0.00         1
         469       0.00      0.00      0.00         0
         470       0.00      0.00      0.00         0
         471       0.00      0.00      0.00         1
         472       0.00      0.00      0.00         0
         473       0.00      0.00      0.00         1
         474       0.00      0.00      0.00         1
         475       0.00      0.00      0.00         1
         476       0.00      0.00      0.00         0
         477       0.00      0.00      0.00         1
         478       0.00      0.00      0.00         0
         479       0.00      0.00      0.00         1
         480       0.00      0.00      0.00         1
         481       0.00      0.00      0.00         1
         482       0.00      0.00      0.00         0
         483       0.00      0.00      0.00         1
         484       0.00      0.00      0.00         1
         485       0.00      0.00      0.00         1
         486       0.00      0.00      0.00         0
         487       0.00      0.00      0.00         0
         488       0.00      0.00      0.00         0
         489       0.00      0.00      0.00         0
         490       0.00      0.00      0.00         1
         491       0.00      0.00      0.00         0
         492       0.00      0.00      0.00         1
         493       0.00      0.00      0.00         1
         494       0.00      0.00      0.00         1
         495       0.00      0.00      0.00         1
         496       0.00      0.00      0.00         1
         497       0.00      0.00      0.00         0
         498       0.00      0.00      0.00         1
         499       0.00      0.00      0.00         0
         500       0.00      0.00      0.00         1
         501       0.00      0.00      0.00         0
         502       0.00      0.00      0.00         1
         503       0.00      0.00      0.00         0
         504       0.00      0.00      0.00         0
         505       0.00      0.00      0.00         1
         506       0.00      0.00      0.00         1
         507       0.00      0.00      0.00         0
         508       0.00      0.00      0.00         1
         509       0.00      0.00      0.00         1
         510       0.00      0.00      0.00         1
         511       0.00      0.00      0.00         0
         512       0.00      0.00      0.00         0
         513       0.00      0.00      0.00         0
         514       0.00      0.00      0.00         0
         515       0.00      0.00      0.00         1
         516       0.00      0.00      0.00         1
         517       0.00      0.00      0.00         1
         518       0.00      0.00      0.00         0
         519       0.00      0.00      0.00         0
         520       0.00      0.00      0.00         0
         521       0.00      0.00      0.00         0
         522       0.00      0.00      0.00         0
         523       0.00      0.00      0.00         0
         524       0.00      0.00      0.00         0
         525       0.00      0.00      0.00         1
         526       0.00      0.00      0.00         1
         527       0.00      0.00      0.00         1
         528       0.00      0.00      0.00         1
         529       0.00      0.00      0.00         0
         530       0.00      0.00      0.00         1
         531       0.00      0.00      0.00         0
         532       0.00      0.00      0.00         1
         533       0.00      0.00      0.00         0
         534       0.00      0.00      0.00         1
         535       0.00      0.00      0.00         0
         536       0.00      0.00      0.00         1
         537       0.00      0.00      0.00         1
         538       0.00      0.00      0.00         0
         539       0.00      0.00      0.00         1
         540       0.00      0.00      0.00         0
         541       0.00      0.00      0.00         1
         542       0.00      0.00      0.00         1
         543       0.00      0.00      0.00         1
         544       0.00      0.00      0.00         0
         545       0.00      0.00      0.00         1
         546       0.00      0.00      0.00         1
         547       0.00      0.00      0.00         0
         548       0.00      0.00      0.00         0
         549       0.00      0.00      0.00         2
         550       0.00      0.00      0.00         0
         551       0.00      0.00      0.00         0
         552       0.00      0.00      0.00         0
         553       0.00      0.00      0.00         0
         554       0.00      0.00      0.00         1
         555       0.00      0.00      0.00         1
         556       0.00      0.00      0.00         1
         557       0.00      0.00      0.00         0
         558       0.00      0.00      0.00         1
         559       0.00      0.00      0.00         0
         560       0.00      0.00      0.00         0
         561       0.00      0.00      0.00         1
         562       0.00      0.00      0.00         1
         563       0.00      0.00      0.00         0
         564       0.00      0.00      0.00         0
         565       0.00      0.00      0.00         1
         566       0.00      0.00      0.00         0
         567       0.00      0.00      0.00         1
         568       0.00      0.00      0.00         1
         569       0.00      0.00      0.00         1
         570       0.00      0.00      0.00         0
         571       0.00      0.00      0.00         1
         572       0.00      0.00      0.00         1
         573       0.00      0.00      0.00         0
         574       0.00      0.00      0.00         1
         575       0.00      0.00      0.00         1
         576       0.00      0.00      0.00         1
         577       0.00      0.00      0.00         0
         578       0.00      0.00      0.00         0
         579       0.00      0.00      0.00         1
         580       0.00      0.00      0.00         0
         581       0.00      0.00      0.00         2
         582       0.00      0.00      0.00         0
         583       0.00      0.00      0.00         1
         584       0.00      0.00      0.00         1
         585       0.00      0.00      0.00         0
         586       0.00      0.00      0.00         0
         587       0.00      0.00      0.00         1
         588       0.00      0.00      0.00         0
         589       0.00      0.00      0.00         0
         590       0.00      0.00      0.00         1
         591       0.00      0.00      0.00         0
         592       0.00      0.00      0.00         1
         593       0.00      0.00      0.00         1
         594       0.00      0.00      0.00         1
         595       0.00      0.00      0.00         1
         596       0.00      0.00      0.00         0
         597       0.00      0.00      0.00         1
         598       0.00      0.00      0.00         1
         599       0.00      0.00      0.00         1
         600       0.00      0.00      0.00         1
         601       0.00      0.00      0.00         0
         602       0.00      0.00      0.00         0
         603       0.00      0.00      0.00         1
         604       0.00      0.00      0.00         0
         605       0.00      0.00      0.00         0
         606       0.00      0.00      0.00         1
         607       0.00      0.00      0.00         0
         608       0.00      0.00      0.00         1
         609       0.00      0.00      0.00         1
         610       0.00      0.00      0.00         1
         611       0.00      0.00      0.00         0
         612       0.00      0.00      0.00         1
         613       0.00      0.00      0.00         1
         614       0.00      0.00      0.00         0
         615       0.00      0.00      0.00         0
         616       0.00      0.00      0.00         1
         617       0.00      0.00      0.00         1
         618       0.00      0.00      0.00         0
         619       0.00      0.00      0.00         0
         620       0.00      0.00      0.00         0
         621       0.00      0.00      0.00         0
         622       0.00      0.00      0.00         1
         623       0.00      0.00      0.00         0
         624       0.00      0.00      0.00         1
         625       0.00      0.00      0.00         1
         626       0.00      0.00      0.00         0
         627       0.00      0.00      0.00         1
         628       0.00      0.00      0.00         0
         629       0.00      0.00      0.00         1
         630       0.00      0.00      0.00         0
         631       0.00      0.00      0.00         1
         632       0.00      0.00      0.00         0
         633       0.00      0.00      0.00         1
         634       0.00      0.00      0.00         1
         635       0.00      0.00      0.00         0
         636       0.00      0.00      0.00         2
         637       0.00      0.00      0.00         1
         638       0.00      0.00      0.00         1
         639       0.00      0.00      0.00         1
         640       0.00      0.00      0.00         1
         641       0.00      0.00      0.00         0
         642       0.00      0.00      0.00         0
         643       0.00      0.00      0.00         1
         644       0.00      0.00      0.00         0
         645       0.00      0.00      0.00         0
         646       0.00      0.00      0.00         0
         647       0.00      0.00      0.00         0
         648       0.00      0.00      0.00         0
         649       0.00      0.00      0.00         0
         650       0.00      0.00      0.00         1
         651       0.00      0.00      0.00         0
         652       0.00      0.00      0.00         0
         653       0.00      0.00      0.00         1
         654       0.00      0.00      0.00         1
         655       0.00      0.00      0.00         0
         656       0.00      0.00      0.00         1
         657       0.00      0.00      0.00         1
         658       0.00      0.00      0.00         1
         659       0.00      0.00      0.00         0
         660       0.00      0.00      0.00         1
         661       0.00      0.00      0.00         1
         662       0.00      0.00      0.00         1
         663       0.00      0.00      0.00         0
         664       0.00      0.00      0.00         0
         665       0.00      0.00      0.00         1
         666       0.00      0.00      0.00         1
         667       0.00      0.00      0.00         0
         668       0.00      0.00      0.00         0
         669       0.00      0.00      0.00         0
         670       0.00      0.00      0.00         0
         671       0.00      0.00      0.00         1
         672       0.00      0.00      0.00         1
         673       0.00      0.00      0.00         0
         674       0.00      0.00      0.00         0
         675       0.00      0.00      0.00         0
         676       0.00      0.00      0.00         1
         677       0.00      0.00      0.00         1
         678       0.00      0.00      0.00         1
         679       0.00      0.00      0.00         0
         680       0.00      0.00      0.00         0
         681       0.00      0.00      0.00         0
         682       0.00      0.00      0.00         0
         683       0.00      0.00      0.00         0
         684       0.00      0.00      0.00         0
         685       0.00      0.00      0.00         1
         686       0.00      0.00      0.00         0
         687       0.00      0.00      0.00         1
         688       0.00      0.00      0.00         1
         689       0.00      0.00      0.00         0
         690       0.00      0.00      0.00         1
         691       0.00      0.00      0.00         1
         692       0.00      0.00      0.00         1
         693       0.00      0.00      0.00         0
         694       0.00      0.00      0.00         1
         695       0.00      0.00      0.00         0
         696       0.00      0.00      0.00         0
         697       0.00      0.00      0.00         0
         698       0.00      0.00      0.00         0
         699       0.00      0.00      0.00         0
         700       0.00      0.00      0.00         1
         701       0.00      0.00      0.00         1
         702       0.00      0.00      0.00         1
         703       0.00      0.00      0.00         0
         704       0.00      0.00      0.00         1
         705       0.00      0.00      0.00         1
         706       0.00      0.00      0.00         0
         707       0.00      0.00      0.00         1
         708       0.00      0.00      0.00         0
         709       0.00      0.00      0.00         1
         710       0.00      0.00      0.00         0
         711       0.00      0.00      0.00         1
         712       0.00      0.00      0.00         1
         713       0.00      0.00      0.00         0
         714       0.00      0.00      0.00         0
         715       0.00      0.00      0.00         0
         716       0.00      0.00      0.00         1
         717       0.00      0.00      0.00         1
         718       0.00      0.00      0.00         0
         719       0.00      0.00      0.00         1
         720       0.00      0.00      0.00         0
         721       0.00      0.00      0.00         1

    accuracy                           0.14       833
   macro avg       0.00      0.01      0.00       833
weighted avg       0.13      0.14      0.13       833
```

Decision Tree

In [372]:
```
model2 = DecisionTreeClassifier()
model2.fit(xtrain,ytrain)
p2 = model2.predict(xtest)
```

In [374]:
```
print(classification_report(ytest.values.ravel(),p2,target_names=['0','1','2','3','4','5','6']))
```

```
              precision    recall  f1-score   support

           0       0.17      0.20      0.18         5
           1       0.17      0.25      0.20         4
           2       0.40      0.40      0.40         5
           3       0.56      0.62      0.59         8
           4       0.75      0.43      0.55         7
           5       0.40      0.25      0.31         8
           6       0.25      0.50      0.33         2

    accuracy                           0.38        39
   macro avg       0.38      0.38      0.37        39
weighted avg       0.43      0.38      0.39        39
```