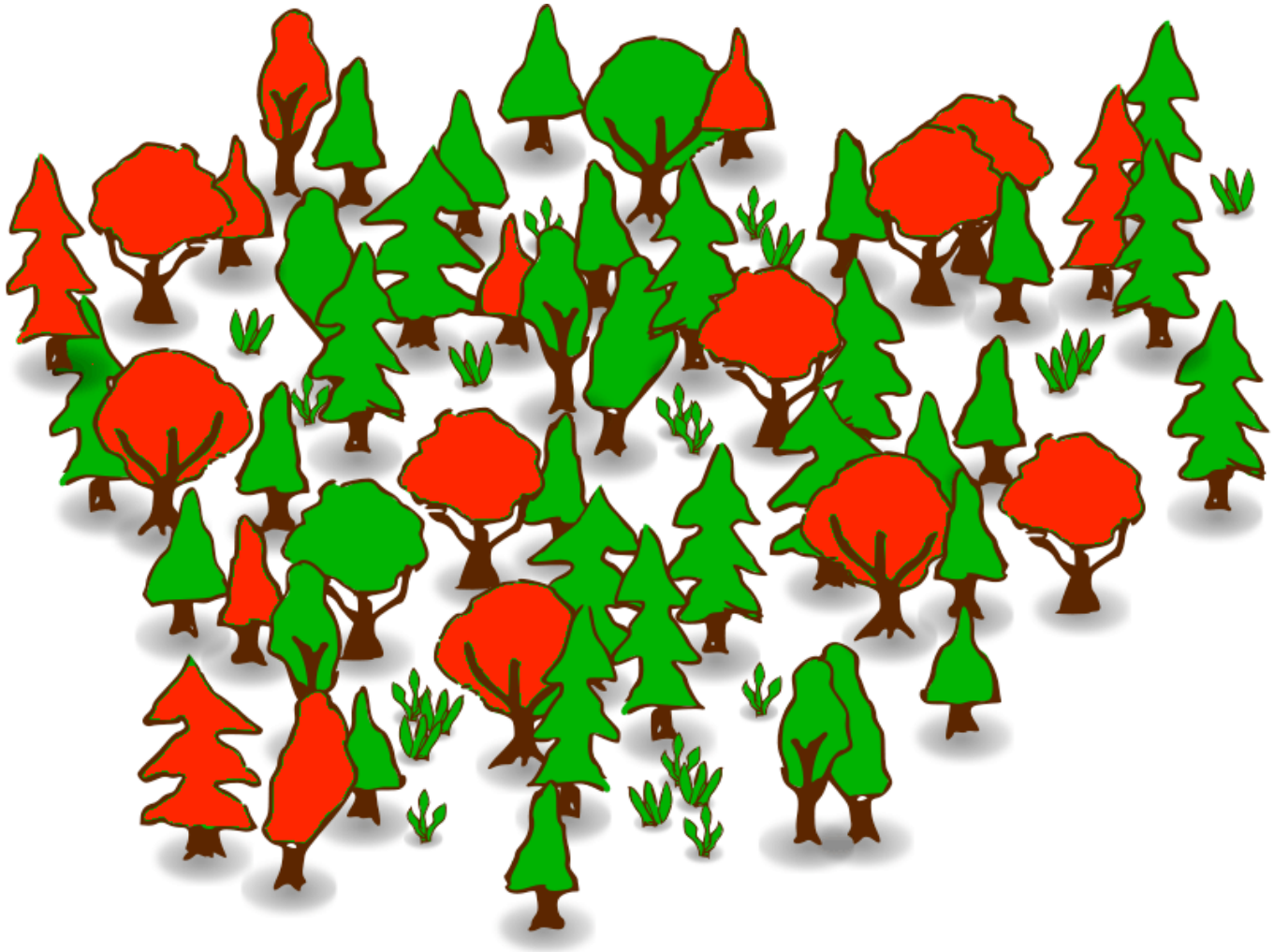
Crimes in the UK

Week 03 – 30.11.2016

Sergiu Soima
Chaoran Chen
Aser Abdelrahman

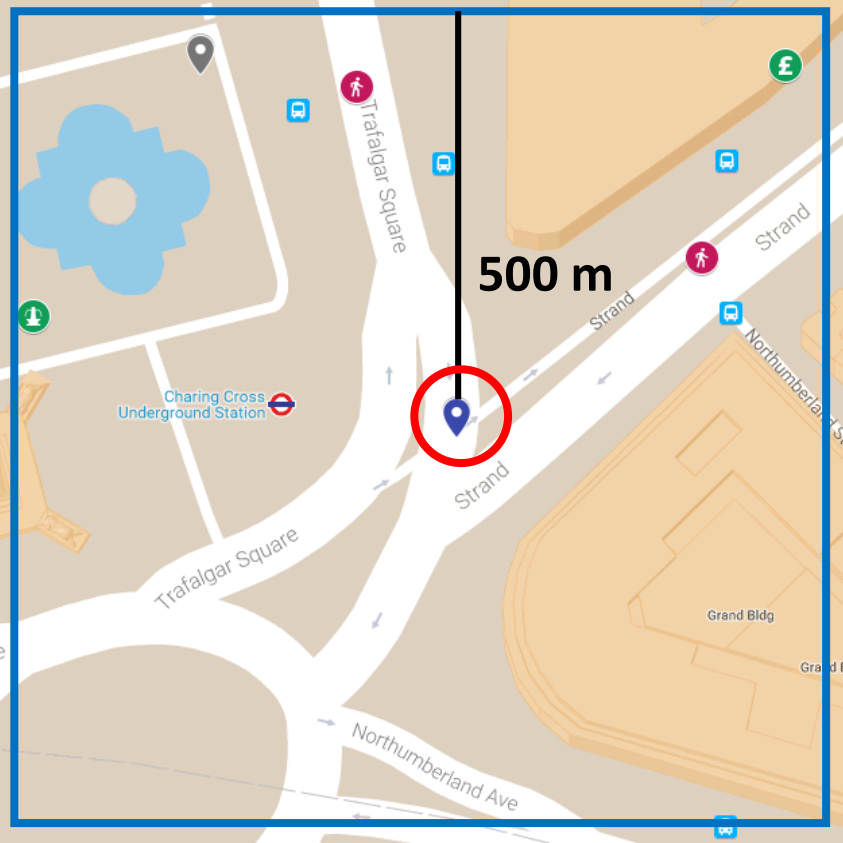


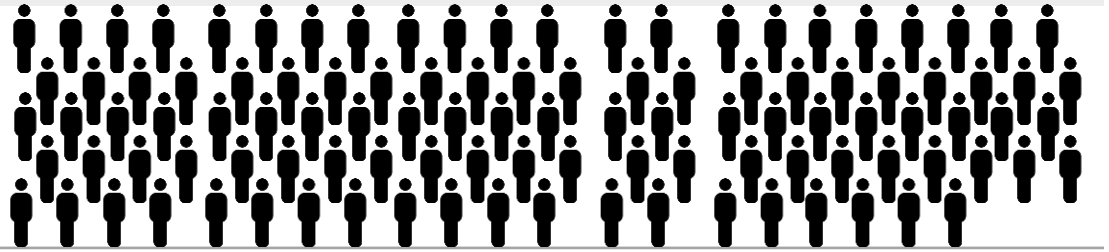
KARTENLEGENDE

Crime 6123122792

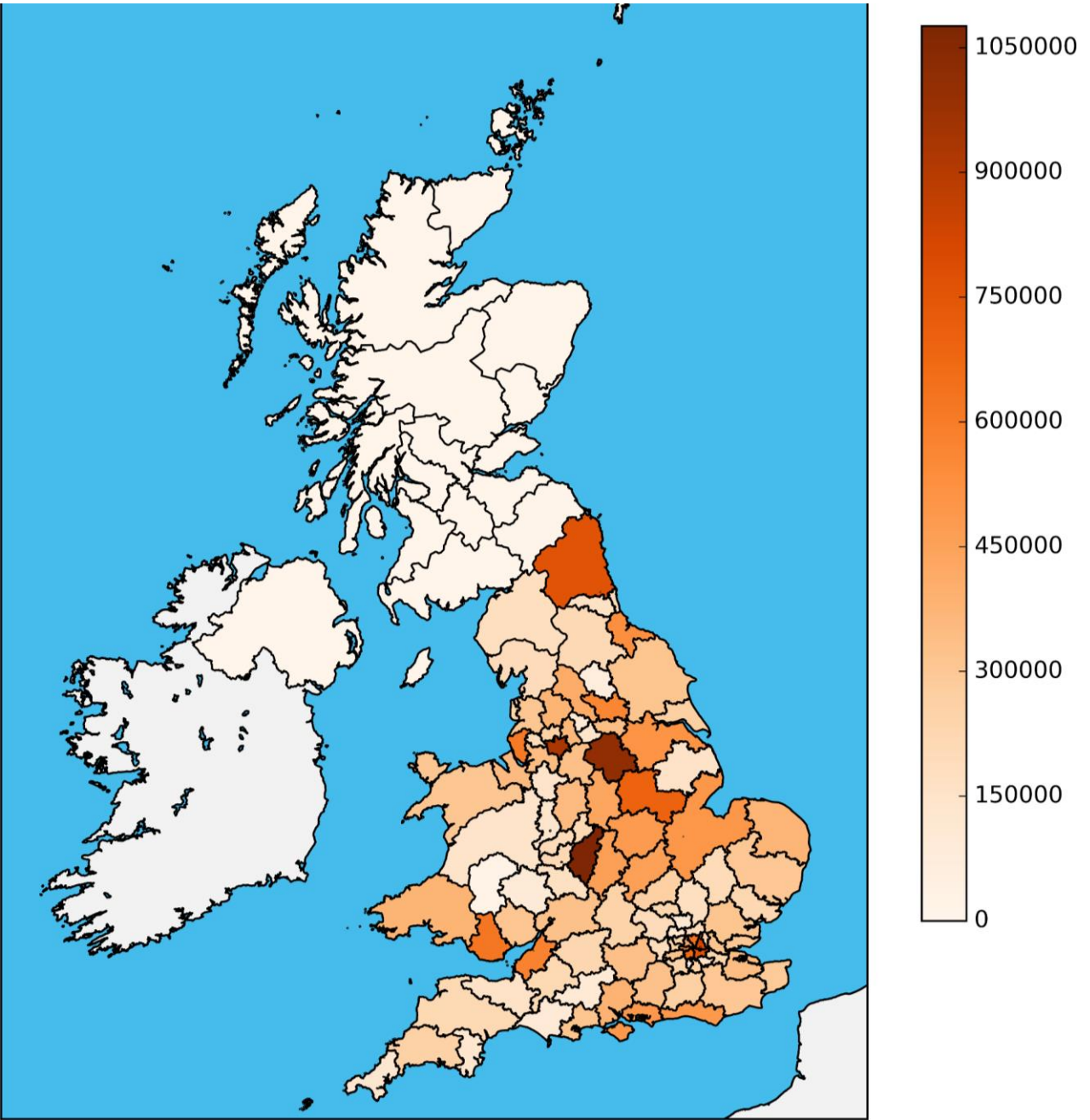
- Crime type: Theft from the person
- Outcome: Defendant found not guilty
- Date: September 2016
- Police force: City of London
- **Latitude / Longitude**
- **Lower Layer Super Output Area (LSOA)**

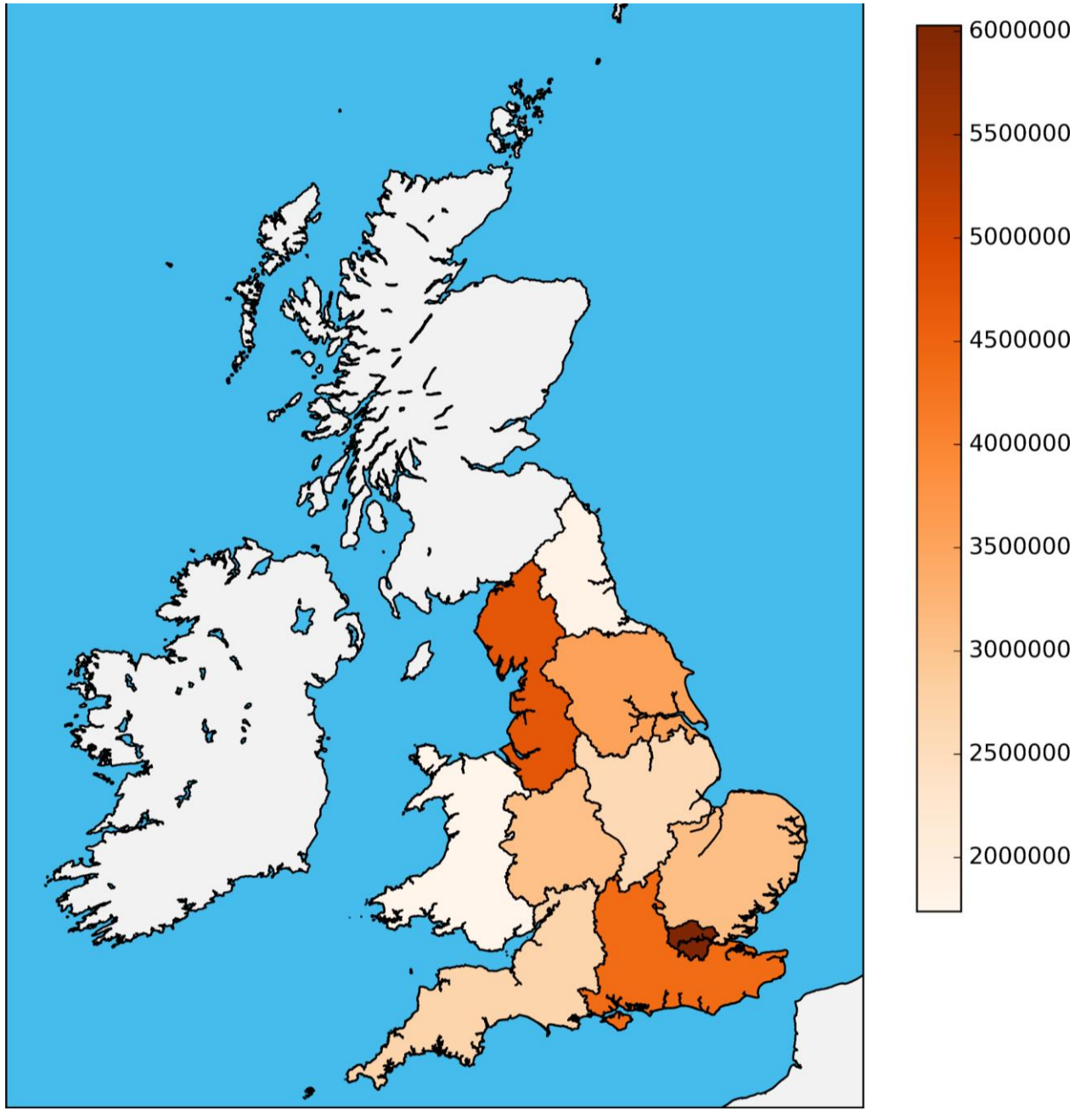
- Number person searches: 2
- Number vehicle searches: 0
- Number fountains: 1
- Number banks 1
- Number Wi-Fi hotspots: 0



Latitude / Longitude	
Lower Layer Super Output Area (LSOA)	
Middle Layer Super Output Areas (MSOA)	
Postcode Area	
Region	

Absolute Number of Crimes





First Approach for Predicting Crime Outcome Types

1	2	3	4	5	6	8	9	...	78
crime_id	**month**	**reported_by**	**longitude**	**latitude**	**lsoa_code**	**crime_type**	**number_poi**	**...**	**number_poi_bank**
29d9fab50e58	01-07-12	City of London Poli	-0.075508	51.514763	E01000005	Criminal dama	113		0
2f0d2ec1570e	01-10-13	City of London Poli	-0.088995	51.512088	E01032739	Other theft	117	...	0
3a1a2883ecc	01-08-14	City of London Poli	-0.102502	51.516165	E01032740	Bicycle theft	105	...	0
3a70044277c	01-06-15	City of London Poli	-0.078393	51.515728	E01000005	Burglary	142	...	0
732de5fe5b99	01-02-12	City of London Poli	-0.080034	51.513389	E01032739	Burglary	147	...	0
f6246f16d4fb	01-05-12	City of London Poli	-0.077578	51.510795	E01032739	Other theft	903	...	4
5f731e098819	01-08-16	City of London Poli	-0.108231	51.513928	E01032740	Shoplifting	108	...	0
6ec7431179e	01-06-14	City of London Poli	-0.109647	51.514887	E01032740	Drugs	115	...	0
07a9dd005b3	01-08-15	City of London Poli	-0.085149	51.517196	E01032739	Other theft	152	...	0

?

79
outcome_type
Investigation comp
Offender sent to p
Investigation comp
Offender given a d
Investigation comp
Suspect charged
Offender sent to p
Offender given a d
Suspect charged

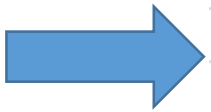
First Approach for Predicting Crime Outcome Types

- Classifier: Random Forest
- Sample subset: Only the city of London (~ 216.000 samples)
- Feature subset: {Crime type, number of POIS, 3 Outcome types of stop& search incidents}
- Library: Scikit-learn Machine Learning in Python
- But, Scikit- learn’s decision trees don’t support categorical features directly…

First Approach for Predicting Crime Outcome Types

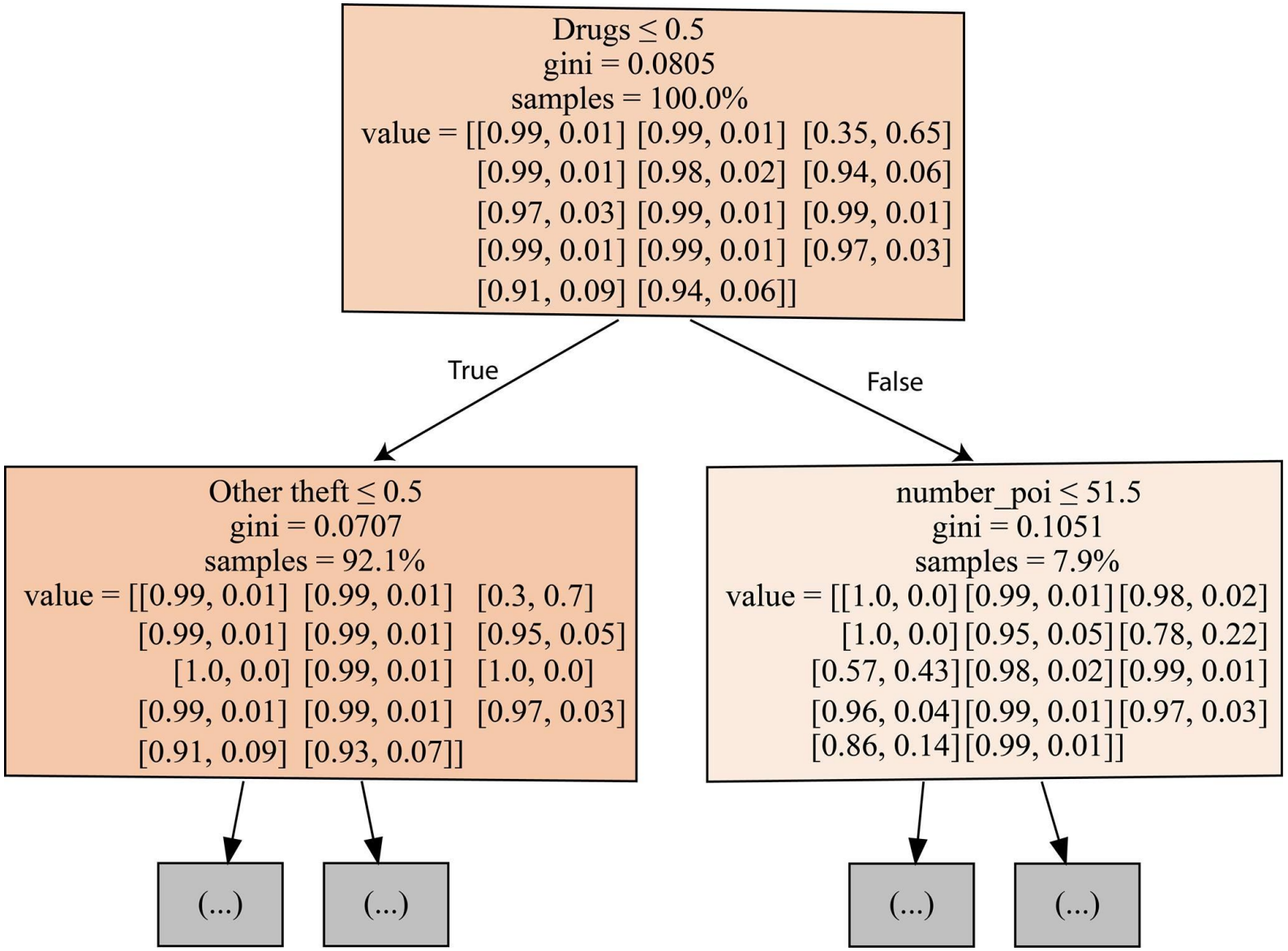
Data preprocessing:

8
crime_type
Criminal dam
Other theft
Bicycle theft
Burglary
Burglary
Other theft
Shoplifting
Drugs
Other theft



1	2	3	4	5	6	11	...	15
Bicycle theft	Burglary	Drugs	Other crime	Other theft	Robbery	Shoplifting	...	Violent crime
0	0	0	0	0	0	0	...	0
0	0	0	0	1	0	0	...	0
0	0	0	0	1	0	0	...	0
0	0	1	0	0	0	0	...	0
1	0	0	0	0	0	0	...	0
0	0	0	0	0	0	0	...	0
0	0	0	0	0	0	1	...	0
0	0	0	0	1	0	0	...	0
0	0	0	0	1	0	0	...	0

First Approach for Predicting Crime Outcome Types



- 10 decision trees in the forest
- cross validation score: 59.0%
- Accuracy of test data: 95.9%
- Weighted F1-score 0.7647

Future Plans:

- Analyze the current random forest
- Improve the current random forest by:
  - Including the new location data
  - Using the whole UK crime dataset
  - Adding new features
- Create the dark map