



CHICAGO DOT CAR CRASHES ANALYSIS

Shayan Abdul Karim Khan

05/08/2023



PROBLEM OVERVIEW



1 FATAL INCIDENT EVERY 15 MINUTES

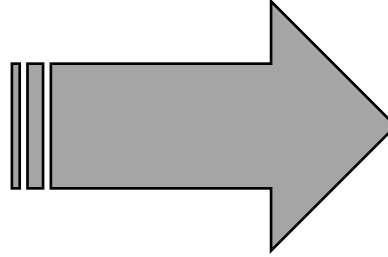
CDOT GOALS



ANALYSIS PROCESS

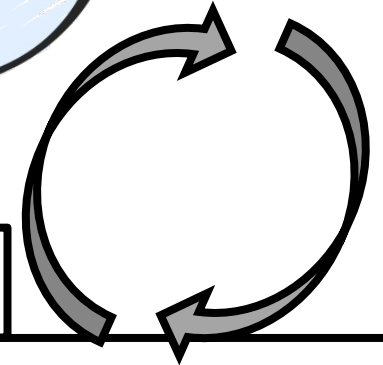



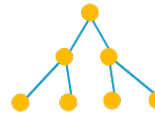
**CHICAGO DOT
DATABASE ON
CRASHES,
VEHICLES AND
PEOPLE**



**ITERATIVE
DATA
CLEANING &
PROCESSING**

MODELLING




Decision Tree


**KNN
Algorithm**

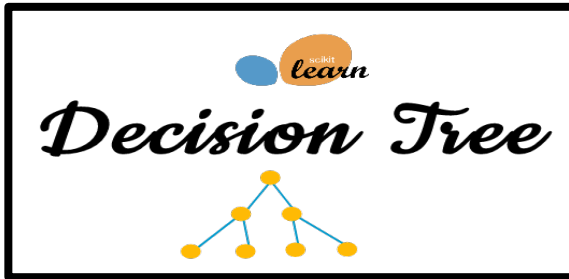


www.educba.com


Random Forest


dmlc
XGBoost

MODELLING RESULTS



ACCURACY: 70%
PRECISION: 70%

4



ACCURACY: 73%
PRECISION: 74%

3



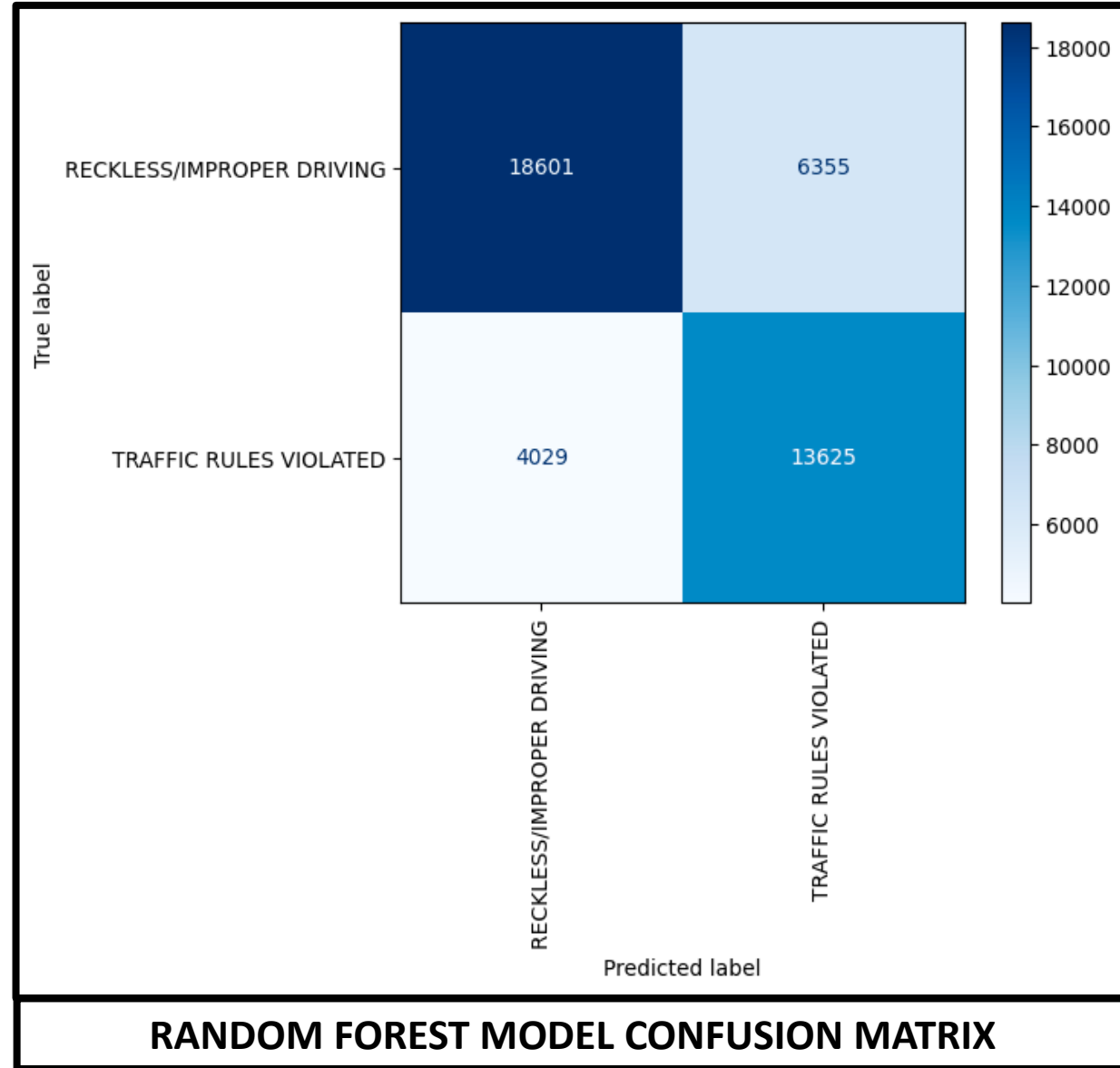
ACCURACY: 74%
PRECISION: 74%

2



ACCURACY: 76%
PRECISION: 76%

1



RANDOM FOREST MODEL CONFUSION MATRIX

TOP FACTORS RESULTS



LATITUDE ★
LONGITUDE ★
INTERSECTION ★



DRIVER ACTION ★
MANEUVER
OCCUPANT CATEGORY



FIRST CONTACT POINT
UNITS
DEVICE CONDITION



WEATHER CONDITION

IMPROVEMENTS

- Higher Granularity
- Increased Model Iterations
- Improve Sampling Methods and fill in gaps in original data
- Use missing factors such as Traffic Volume and Congestion Percentages

NEXT STEPS - INSIGHTS

- Research the trends tied with location factors
- Explore how driver actions impact crashes
- Develop mobility engineering plans and awareness campaigns
- Improve performance of model

QUESTIONS?

CHICAGO DOT CAR CRASHES ANALYSIS



05/08/2023

BACKUP



	Feature	Score
15	LATITUDE	0.575710
16	LONGITUDE	0.571104
26	DRIVER_ACTION	0.152321
5	INTERSECTION_RELATED_I	0.081477
18	MANEUVER	0.072786
19	FIRST_CONTACT_POINT	0.053739
7	NUM_UNITS	0.045599
2	DEVICE_CONDITION	0.041522
3	WEATHER_CONDITION	0.041163
28	PHYSICAL_CONDITION	0.033478

```
#check new value counts
```

```
cln_people['DRIVER_ACTION'].value_counts()
```

NONE	395225
PASSENGER	185993
UNKNOWN	152429
OTHER	84170
FAILED TO YIELD	78998
FOLLOWED TOO CLOSELY	50961
IMPROPER TURN	22215
IMPROPER BACKING	22027
IMPROPER LANE CHANGE	19951
IMPROPER PASSING	13647
DISREGARDED CONTROL DEVICES	13157
TOO FAST FOR CONDITIONS	12684
WRONG WAY/SIDE	2716
IMPROPER PARKING	2603
OVERCORRECTED	1513
CELL PHONE USE OTHER THAN TEXTING	1343
EVADING POLICE VEHICLE	1148
EMERGENCY VEHICLE ON CALL	926
TEXTING	351
STOPPED SCHOOL BUS	121
LICENSE RESTRICTIONS	44

Name: DRIVER_ACTION, dtype: int64

```
#check value counts
```

```
cln_veh['OCCUPANT_CAT'].value_counts()
```

1-2	1102688
0	180457
3-5	54378
6-9	1979
10+	306

Name: OCCUPANT_CAT, dtype: int64