

1s

```
df_train=pd.read_csv(url1,parse_dates=['DateTime'])
df_train.head()
```



	DateTime	Junction	Vehicles	ID
0	2015-11-01 00:00:00	1	15	20151101001
1	2015-11-01 01:00:00	1	13	20151101011
2	2015-11-01 02:00:00	1	10	20151101021
3	2015-11-01 03:00:00	1	7	20151101031
4	2015-11-01 04:00:00	1	9	20151101041



```
df_test=pd.read_csv(url2,parse_dates=['DateTime'])
df_test.head()
```



	DateTime	Junction	ID
0	2017-07-01 00:00:00	1	20170701001
1	2017-07-01 01:00:00	1	20170701011
2	2017-07-01 02:00:00	1	20170701021
3	2017-07-01 03:00:00	1	20170701031
4	2017-07-01 04:00:00	1	20170701041



✓ [5] df\_train.shape

0s

(48120, 4)

✓ [6] df\_test.shape

0s

(11808, 3)

✓ [7] df\_train.isnull().sum()

0s

```
DateTime    0
Junction    0
Vehicles    0
ID          0
dtype: int64
```

✓ df\_test.isnull().sum()

0s

```
DateTime    0
Junction    0
ID          0
dtype: int64
```

▶ Xy\_train= gen\_lag\_features(train)  
Xy\_train

↑ ↓ ↻ ⌂ ⚙ 📄 🗑

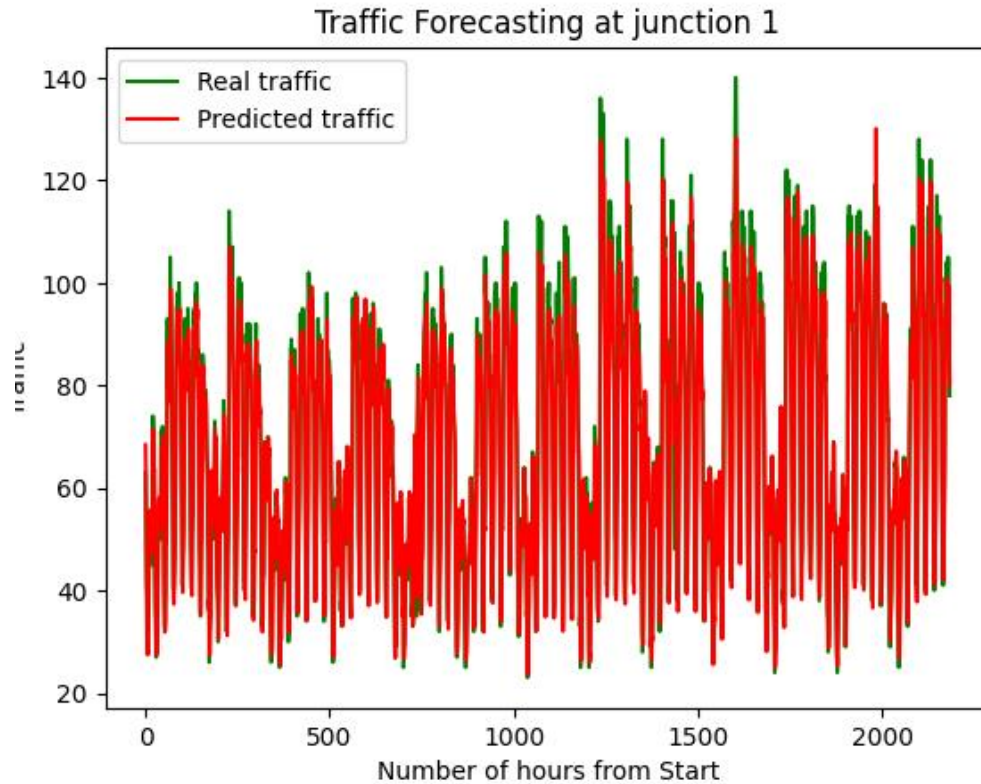


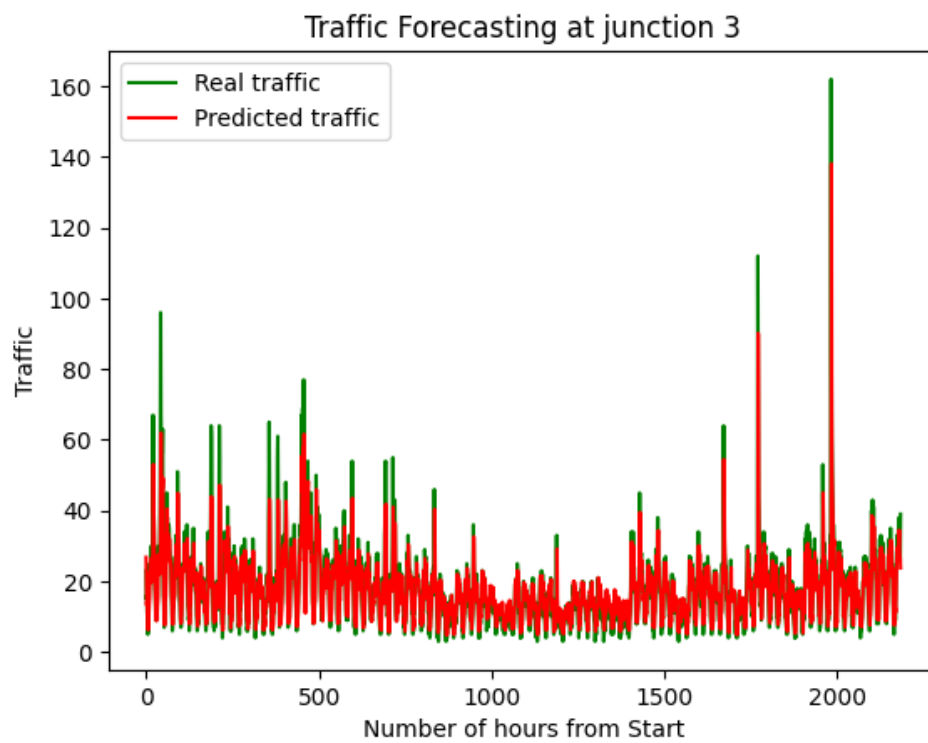
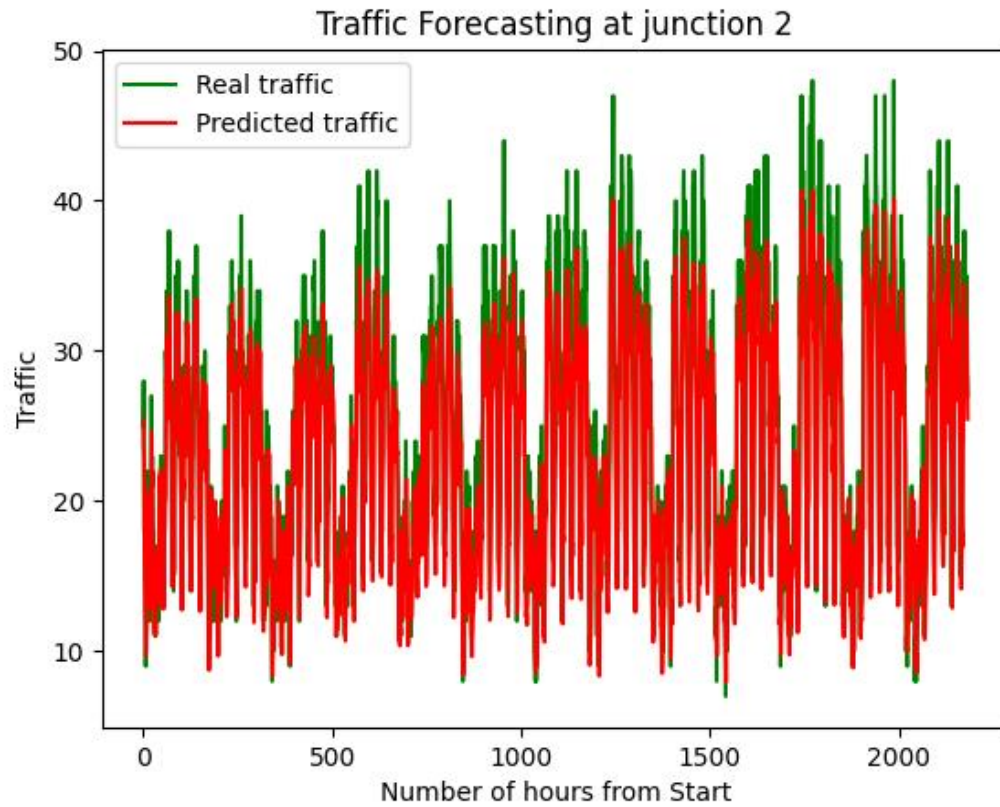
	Junction 1 (H-1)	Junction 2 (H-1)	Junction 3 (H-1)	Junction 4 (H-1)	Junction 1 (H)	Junction 2 (H)	Junction 3 (H)	Junction 4 (H)
DateTime								
2015-11-01 01:00:00	15.0	6.0	9.0	36.0	13.0	6.0	7.0	36.0
2015-11-01 02:00:00	13.0	6.0	7.0	36.0	10.0	5.0	5.0	36.0
2015-11-01 03:00:00	10.0	5.0	5.0	36.0	7.0	6.0	1.0	36.0
2015-11-01 04:00:00	7.0	6.0	1.0	36.0	9.0	7.0	2.0	36.0
2015-11-01 05:00:00	9.0	7.0	2.0	36.0	6.0	2.0	2.0	36.0
...	...	...	...	...	...	...	...	...
2017-06-30 19:00:00	95.0	34.0	38.0	17.0	105.0	34.0	33.0	11.0
2017-06-30 20:00:00	105.0	34.0	33.0	11.0	96.0	35.0	31.0	30.0
2017-06-30 21:00:00	96.0	35.0	31.0	30.0	90.0	31.0	28.0	16.0
2017-06-30 22:00:00	90.0	31.0	28.0	16.0	84.0	29.0	26.0	22.0
2017-06-30 23:00:00	84.0	29.0	26.0	22.0	78.0	27.0	39.0	12.0

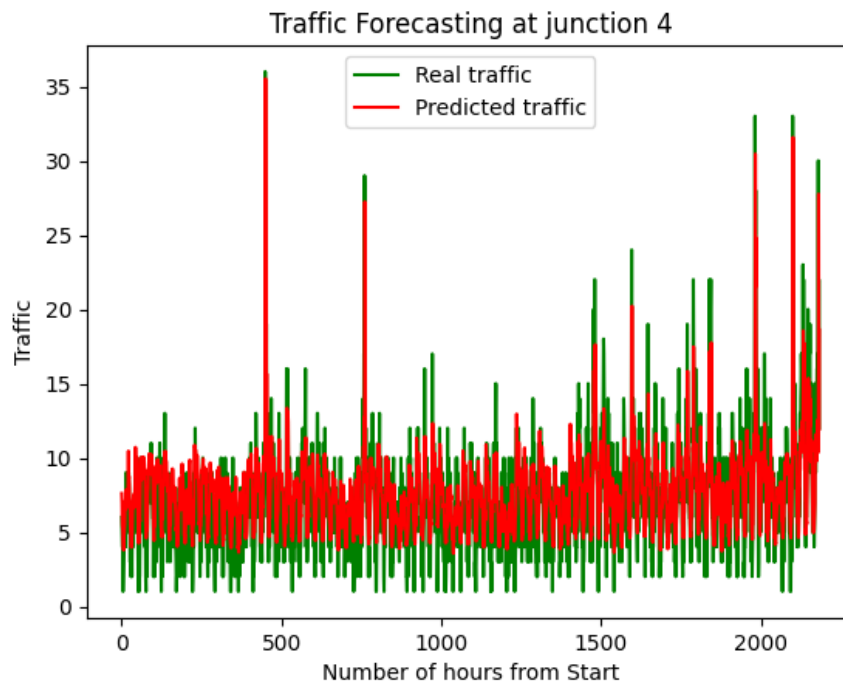
14591 rows x 8 columns

	Junction 1 (H)	Junction 2 (H)	Junction 3 (H)	Junction 4 (H)
DateTime				
2015-11-01 01:00:00	0.052980	0.106383	0.033520	1.000000
2015-11-01 02:00:00	0.033113	0.085106	0.022346	1.000000
2015-11-01 03:00:00	0.013245	0.106383	0.000000	1.000000
2015-11-01 04:00:00	0.026490	0.127660	0.005587	1.000000
2015-11-01 05:00:00	0.006623	0.021277	0.005587	1.000000
...	...	...	...	...
2017-03-31 19:00:00	0.496689	0.531915	0.156425	0.200000
2017-03-31 20:00:00	0.483444	0.638298	0.156425	0.200000
2017-03-31 21:00:00	0.403974	0.574468	0.150838	0.228571
2017-03-31 22:00:00	0.423841	0.553191	0.162011	0.142857
2017-03-31 23:00:00	0.417219	0.553191	0.162011	0.142857

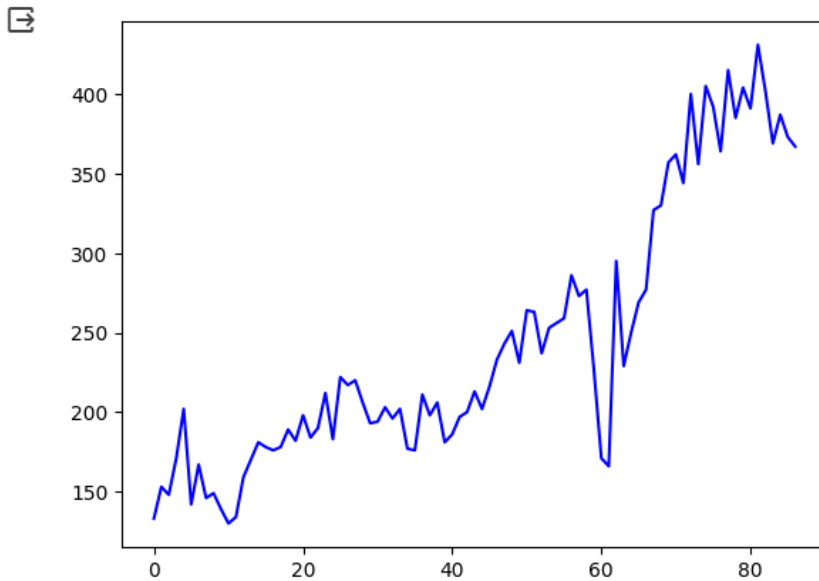
12407 rows x 4 columns







```
import matplotlib.pyplot as plt
for i in range(len(week)):
    plt.plot(week[i], 'blue')
plt.show()
```



```
import matplotlib.pyplot as plt
plt.figure()
plt.plot(jun[0], 'yellow')
plt.show()
plt.plot(jun[1], 'red')
plt.show()
plt.plot(jun[2], 'green')
plt.show()
plt.plot(jun[3], 'blue')
plt.show()
plt.plot(jun[4], 'red')
plt.show()
```

