Predicting traffic flow in a city using induction loop sensors

Challenge Provider: Porto Digital / Municipality of Porto

Context

According to the European Environment Agency, passenger cars are a significant polluter, accounting for 60.7% of total CO2 emissions from road transport in Europe. Without a doubt, congested cities are a significant contributor to this.
The city of Porto has been counting cars at several locations with Inductive Loop sensors since 2015. With this historical data, it should be possible to predict the traffic flow and especially intense periods of traffic.
By predicting the traffic flow in the city and understanding the factors that influence the traffic, it's possible to take action into reducing the traffic.

Goals

On the one hand, the goal of this challenge is to create an explainable prediction of when heavy flow and congestions in the city will happen. On the other hand, it shows that it is possible to extract useful insight regarding the traffic by using sensors in several locations. By achieving this, it is possible to create tools for the city to manage the traffic much better by achieving this, thus reducing the CO2 emissions.

Outcome

The outcome of this challenge is three-fold:

- Create a predictive model for traffic flow in the city of Porto for different periods of the day from sensor data;
- Explain which factors affect the traffic flow;
- Show any relationship and insights that exist between the different traffic flow in the city.

For this challenge, you **must use sensor data from induction loops,** but you can also use other data.

Available Resources

As a reminder, all the data resources can be found here: https://bit.ly/wdl-data. You can also use any open, free and legally available data - Even if it is for another city. In this case, the main source of data should be induction loop sensors.

Sensor data from induction loops
The sensor data is periodically recorded with the number of cars that pass it.

Data Entities
A small dataset with all the locations of the sensors.

POI
A list of Points of Interests in the city of Porto is provided as well.

Other datasets have been available by the city hall, such air quality, in the case it can be useful.

Submissions

Deadline: 01 - 05 - 2021 @ 14h00 GMT + 1

Don't forget that you will need to deliver the report **using the template provided** (see below) and a 1-minute summary.

Submission template: http://bit.ly/wdl-template

Tips

- **This is pure sensor data**. Naturally, it will be noisy and imperfect. Consider it when working with the datasets.
- Do state-of-the-art research. There might be already a lot of interesting things done before;
- Try to fill in the template from start to finish with a straightforward dummy solution first and iterate afterwards;
- You can use other data sources, such as weather, which can be very useful;
- We don't define which period of the day for you to predict on purpose. We want to tell us which is the most useful in this case;
- If possible, don't forget to explain the predictions of your model.