

Big Data Analytics of Road Traffic Dataset Using Machine Learning Techniques

Introduction:

The project aims at improving traffic conditions and safety using state of the art big data processing engines (e.g. Hadoop, Spark and Flink), to perform analysis on real road traffic data. The traffic data comprises of data taken from various sensors placed on Stockholm highways. These sensors measure various simple but useful metrics such as, the number of cars passing and average speed of the cars etc.

The work will involve techniques like batch processing, stream processing, machine learning and traffic flow theory in order to clean, represent and process the traffic data for useful analysis.

Use-cases:

Some interesting use-cases to give more idea about the project:

- *Anomaly Detection*
Detecting something unusual in the behaviour of traffic flow falls in this category. For example: natural disaster, accident, marathon or construction can cause a change in the traffic flow.
- *Data fusion*
Traffic data can be affected by various conditions such as, weather, holidays, time and location. How this extra environment related information can help us making better analysis of the data would be an interesting use-case to study.
- *Traffic Safety*
Traffic data can be analysed for improving the safety of passengers. For example, detecting end of queues in case of a traffic jam for warning other drivers.
- *Prediction*
Traffic analysis can be based on learning from historical data and using it to predict what would be the traffic condition in short or long term.

Contacts:

- Ahmad Al-Shishtawy, PhD, Senior Researcher. Email: ahmad.al-shishtawy@ri-se
- Zainab Abbas, PhD Student. Email: zainabab@kth.se