A tutorial on network-wide multi-horizon traffic forecasting with deep learning

G. Buroni [1], G. Bontempi [1], K. Determe [2]

[1] Machine Learning Group, ULB, Bruxelles, Belgium

[2] Bruxelles Mobilite, Bruxelles, Belgium

4th International Workshop on Big Mobility Data Analytics, March 23, 2021

MOBIAID Project

Brussels MOBI-AID (Brussels MOBIlity-Advanced Indicators Dashboard) aims at designing and set up a performance monitoring system by means of advanced mobility indicators that allow to:

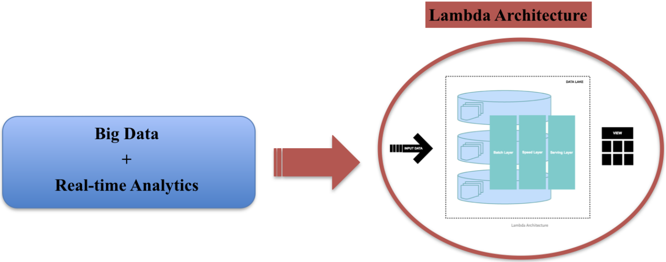
- better understand mobility dynamics in Brussels Region,
- support local authorities in designing suitable and sustainable policies in short and long-term.

VIApass for OBU Data

- Since the 1st of April 2016 Viapass Tracking System collects GPS data about Heavy Good Vehicles (HGV) driving in Belgium by means of On Board Unit (OBU);
- OBUs send information every 30 seconds about:
  - Identifier
  - Timestamp
  - GPS Position
  - Speed
- Bruxelles Mobility Administration collects Viapass data to better understand the traffic dynamics in Bruxelles Capital Region (BCR).

Lambda Architecture for OBU Data

A scalable architecture for distributed storage and processing is required due to **large volume** and **streaming nature** of the OBU data (*Buroni et al., 2018*).



- **Scalable**
- **Fault tolerant**
- **low latency**
- **near real-time**

Multivariate and Multi-horizon Traffic Forecasting (*Bontempi, Taieb, and Le Borgne, 2012*)

GPS sensors allow to monitor the traffic conditions of the entire road network in near real-time.

NETWORK-WIDE FORECASTING

The aim of traffic forecasting is to predict future traffic conditions of the entire transportation networks given a sequence of historical traffic observations.

MULTI-HORIZON FORECASTING

The aim of traffic forecasting is also to predict traffic conditions for multiple steps into the future.

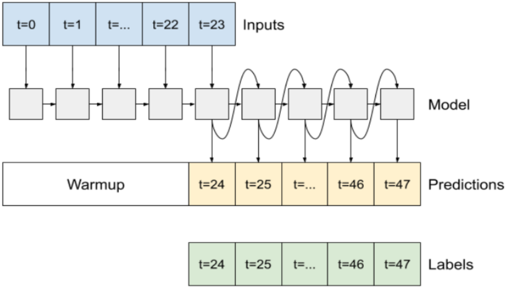
WHY Deep Learning for Traffic Forecasting?

Deep Learning (DL) models are particularly suitable for the task of network-wide multi-horizon traffic forecasting (*Wang, Cao, and Yu, 2020*):

- natively support sequence input data (sequence of traffic observations);
- directly support multiple input sequences for multivariate forecasting (multiple street segments);
- map input sequences (past traffic observations) directly to output sequences (multi-horizon predictions).

Iterated VS Direct Approaches [1] (Lim et al., 2019)

Iterated Model

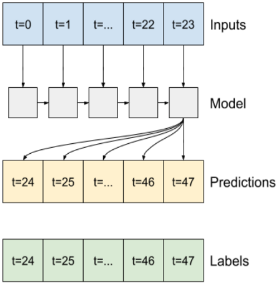


One-step-ahead prediction model where each output is recursively fed into the model.

[1] *https://www.tensorflow.org/tutorials/structured_data/time_series*

Iterated VS Direct Approaches

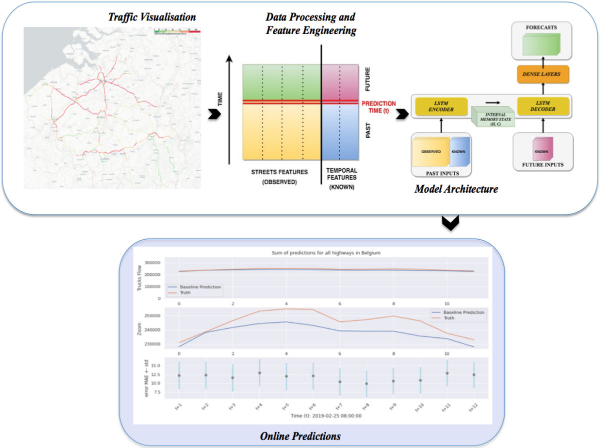
Direct Model



Model is trained to explicitly generate forecasts for multiple horizons in a single step.

Tutorial Outline: Direct LSTM-based Encoder Decoder

Notebook on Kaggle - *here*

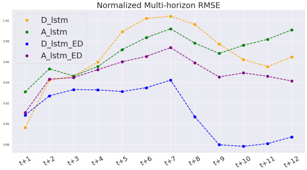


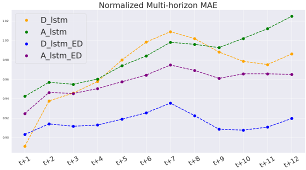
Model Comparison

- Seasonal Persistence Model - Baseline (SW)
- Simple LSTM Model - Direct Approach (D_lstm)
- Simple LSTM Model - Iterated Approach (A_lstm)
- LSTM-based Encoder-Decoder - Iterated Approach (A_lstm_ED)

Results & Conclusion

Average value NRMSE and NMAE for different forecast horizons.





Model	*avgRMSE*	*avgMAE*
Baseline	11.67	7.25
D_lstm	11.28	7.03
A_lstm	11.30	7.12
D_lstm_ED	**10.66**	**6.63**
A_lstm_ED	11.08	6.93

Table 1: ***Average values of RMSE and MAE.***

References

- Bontempi, Gianluca, Souhaib Ben Taieb, and Yann-Aël Le Borgne (2012). "Machine learning strategies for time series forecasting". In: *European business intelligence summer school*. Springer, pp. 62–77.
- Buroni, Giovanni et al. (2018). "On-Board-Unit Data: A Big Data Platform for Scalable storage and Processing". In: *2018 4th International Conference on Cloud Computing Technologies and Applications (Cloudtech)*. IEEE, pp. 1–5.
- Lim, Bryan et al. (2019). "Temporal fusion transformers for interpretable multi-horizon time series forecasting". In: *arXiv preprint arXiv:1912.09363*.
- Wang, Senzhang, Jiannong Cao, and Philip Yu (2020). "Deep learning for spatio-temporal data mining: A survey". In: *IEEE Transactions on Knowledge and Data Engineering*.