POLITECNICO MILANO 1863

July 2024

Traffic Forecasting

Network Data Analysis Lab
Professor Musumeci Francesco
Project 8, Group G

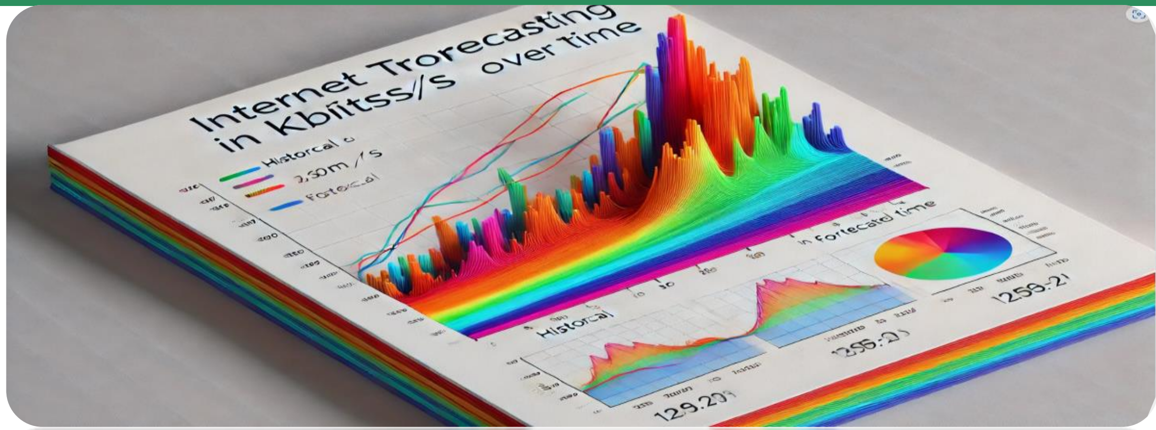
Directed By
Mohammad Amirifard
Dorsa Moadeli
Sajjad Salari

SCAN ME

TABLE OF CONTENTS

1. GENERAL OVERVIEW

1_GEANT Dataset of 1 Month with 15ms Intervals captured in html files



2_Traffic forecasting of dataset with Kbits/s metric

3_Considering on MAE, MSE, MAPE, Training time

4_Compare Local models vs a global model

2. Exploratory Data Analysis

1 Generate Features

Generate a data frame Including all traffics captured during 1 month with 15ms interval

2 Float64 to Float32

Reduce training computations by decreasing bits of numbers to store

3 Remove NaNs

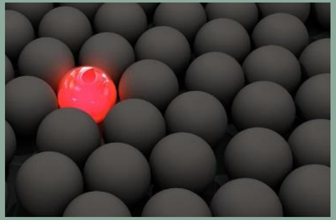
Remove those traffics including more than 25% NaN values

4 Filling NaNs

Fill those traffics including less than 25% NaN values by forward filling method

2. Exploratory Data Analysis

5
Find Anomalies



- Set a window size W and a threshold
- Detector will compare a value to the median or mean of its preceding time window. If it greater or lower, it is recognized as Anomaly

6
Replace Anomalies

First Set Anomalies to NaN and then fill them with forward filling method

7
Save data frames

Save two data frames
1_clean_traffic_dataframe_with_outliers
2_clean_traffic_dataframe_without_outliers

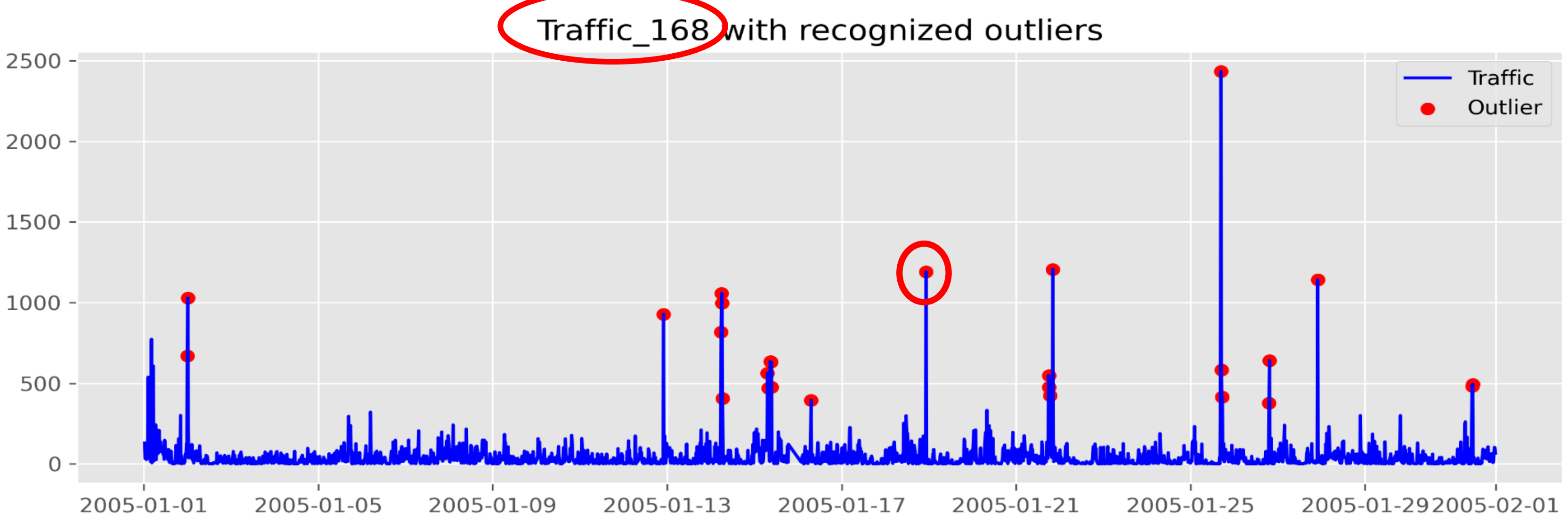
Traffic_5 and Traffic_73 will be selected to work on

2. Exploratory Data Analysis

Let's see the generated data frame
Including features

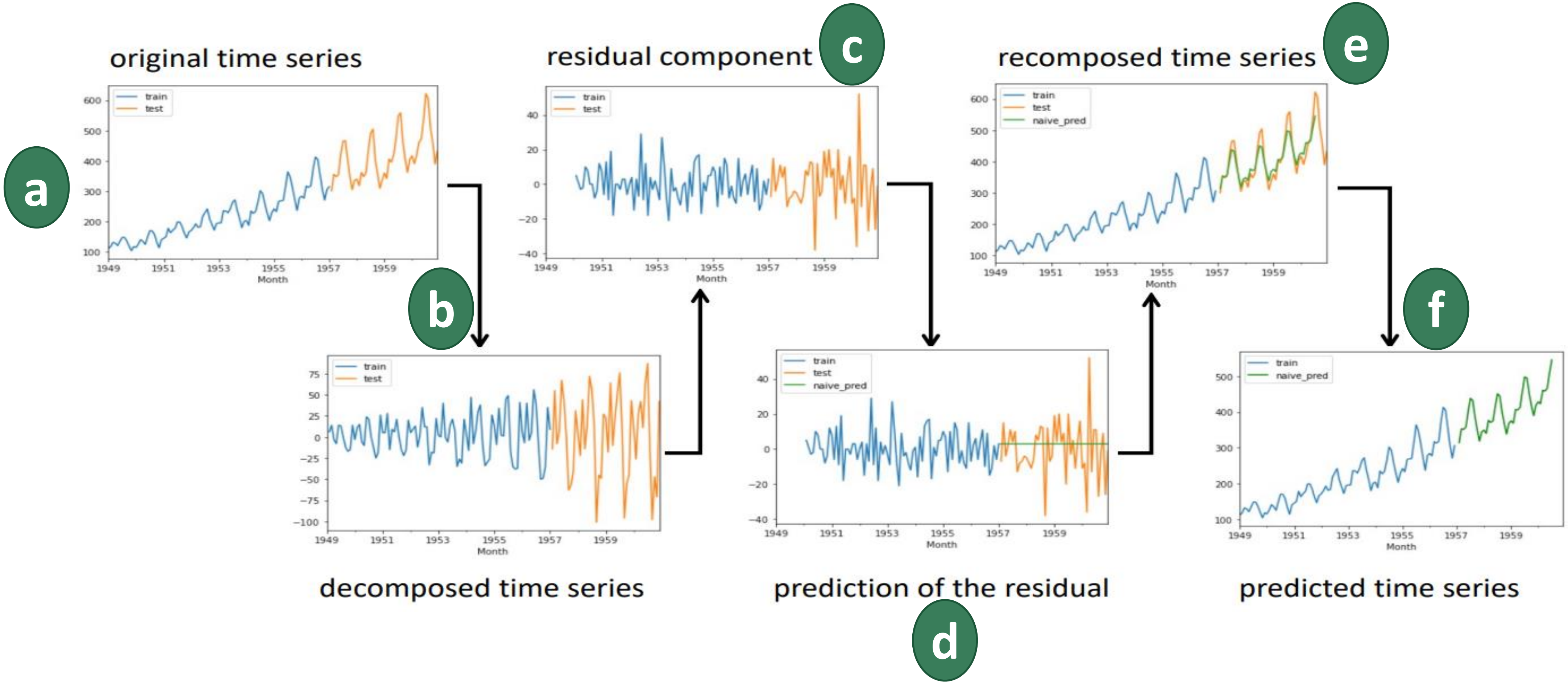
index	src_id	dst_id	2005-01-01-00-30	2005-01-01-00-45	...	2005-01-31-23-45
traffic_0	12	12	396708.2188	338875.5625	...	310534.2813
traffic_1	12	13	28093.75195	28485.3418	...	30300.66602
traffic_2	12	19	16916.56055	11769.3418	...	10089.44922
traffic_3	12	23	3662.808838	1950.506714	...	119.0400009
traffic_4	12	8	6552.933105	5445.911133	...	1205.466675
traffic_5	12	18	3224.915527	3609.431152	...	85431.8125
traffic_6	12	4	124992.7188	120400.4531	...	34571.57422
traffic_7	12	1	14946.89746	15535.51074	...	8623.040039
traffic_8	12	5	22715.86719	10914.40918	...	5257.582031
traffic_9	12	3	19470.03516	13298.21289	...	14440.66699

Let's see anomalies

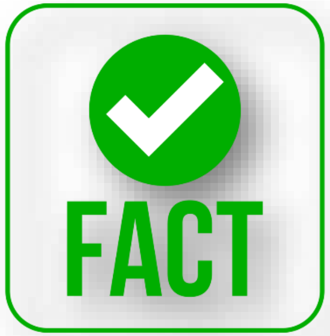


3. Predicting By Residual traffics

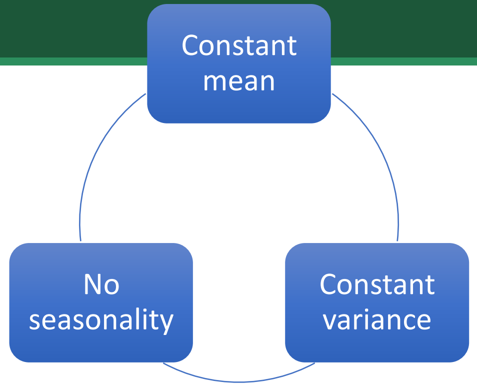
❖ **How can predict traffic by residuals???**



3. Predicting By Residual traffics b



Satationary



❖ If a time series (traffic) is stationary

❖ It is predictable

❖ If a time series (traffic) is not stationary

❖ It is not predictable well

Let's make them Stationary

3. Predicting By Residual traffics b

❖ How to check stationary???

1 Augmented Dickey-Fuller

2 KPSS

•**Null Hypothesis (H0)**: The series has a unit root, meaning it is non-stationary. It contains a structure that may generate trends or random walks.

• **Alternative Hypothesis (H1)**: The series does not have a unit root, meaning it is stationary.

•**Null Hypothesis (H0)**: The time series is stationary around a deterministic trend (or level, depending on the version of the test used).

•**Alternative Hypothesis (H1)**: The time series has a unit root, meaning it is non-stationary and exhibits a stochastic trend.

$$\text{Decision} = \begin{cases} \text{"Stationary"} & \text{if } p\text{-value} \leq 0.05 \\ \text{"Non-stationary"} & \text{if } p\text{-value} > 0.05 \end{cases}$$

$$\text{Decision} = \begin{cases} \text{"Non-stationary"} & \text{if } p\text{-value} \leq 0.05 \\ \text{"Stationary"} & \text{if } p\text{-value} > 0.05 \end{cases}$$

3. Predicting By Residual traffics b

❖ How to check stationary???

1 Augmented Dickey-Fuller

2 KPSS

Obtained Stationary Data frame

index	src_id	dst_id	ADF_test	KPSS_test	Stationary_status
traffic_0	12	12	1	0	0
traffic_1	12	13	1	1	1
traffic_2	12	19	1	0	0
traffic_3	12	23	1	0	0
traffic_4	12	8	1	0	0
traffic_5	12	18	1	0	0
traffic_6	12	4	1	0	0
traffic_7	12	1	1	0	0
traffic_8	12	5	1	0	0
traffic_9	12	3	1	0	0
traffic_10	12	22	1	1	1
traffic_11	12	7	1	0	0
traffic_12	12	2	1	0	0
...	...	...	...	...	...
traffic_438	15	21	1	1	1

3. Predicting By Residual traffics b

How to make non-staionary time series stationary???

Multiplication
Not summation

Decomposing time series to its components

Time Series

Trend Component ⊗ **Seasonal** Component ⊗ **Residual** Component

Overall, persistent, long-term change

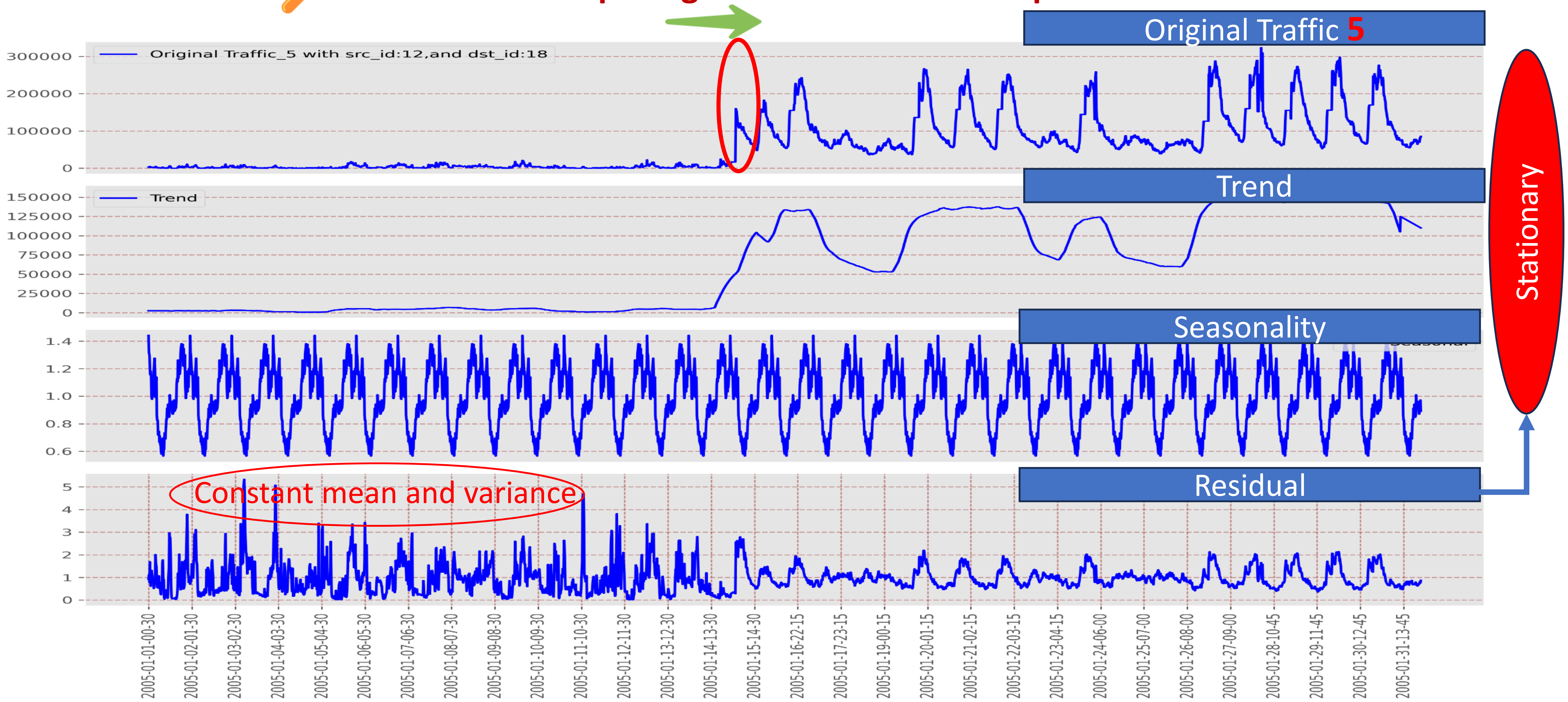
Regular periodic fluctuations

Residual fluctuations

(hopefully stationary)

3. Predicting By Residual traffics c

See one decomposing time series to its component



3. Predicting By Residual traffics d

Set Up models for predicting

Simple Task

Local models

1 RNN

2 CONV_LSTM

Advanced Task

Global model

1 Transfomer

3. Predicting By Residual traffics d

1. Callbacks
2. Hyperparameters
3. Splitting dataset
4. Create Sequences
5. Set RNN/LSTM configuration
6. Train
7. Test Performance
8. Tuning
9. Predicting

Local models

Steps to Create both Local RNN and LSTM models

RNN

SimpleRNN → Dropout → SimpleRNN → SimpleRNN → SimpleRNN → Dense

LSTM

Bi-directional LSTM → Conv1D → Conv1D → GlobalMaxPooling

3. Predicting By Residual traffics d

- Concatenating
- Normalization
- Create Sequences
- Set-up transformer
- Train
- Test Performance
- Tuning
- Predicting

Global model

Steps to Create Global model Using transformers

Dense → Positional Encoding → Transformer blocks → Global Average Pooling1D → Dense

MultiHeadAttention → LayerNormalization → Dense → Dense

After Setting up models
We ran them on

1_clean_traffic_dataframe_with_outliers

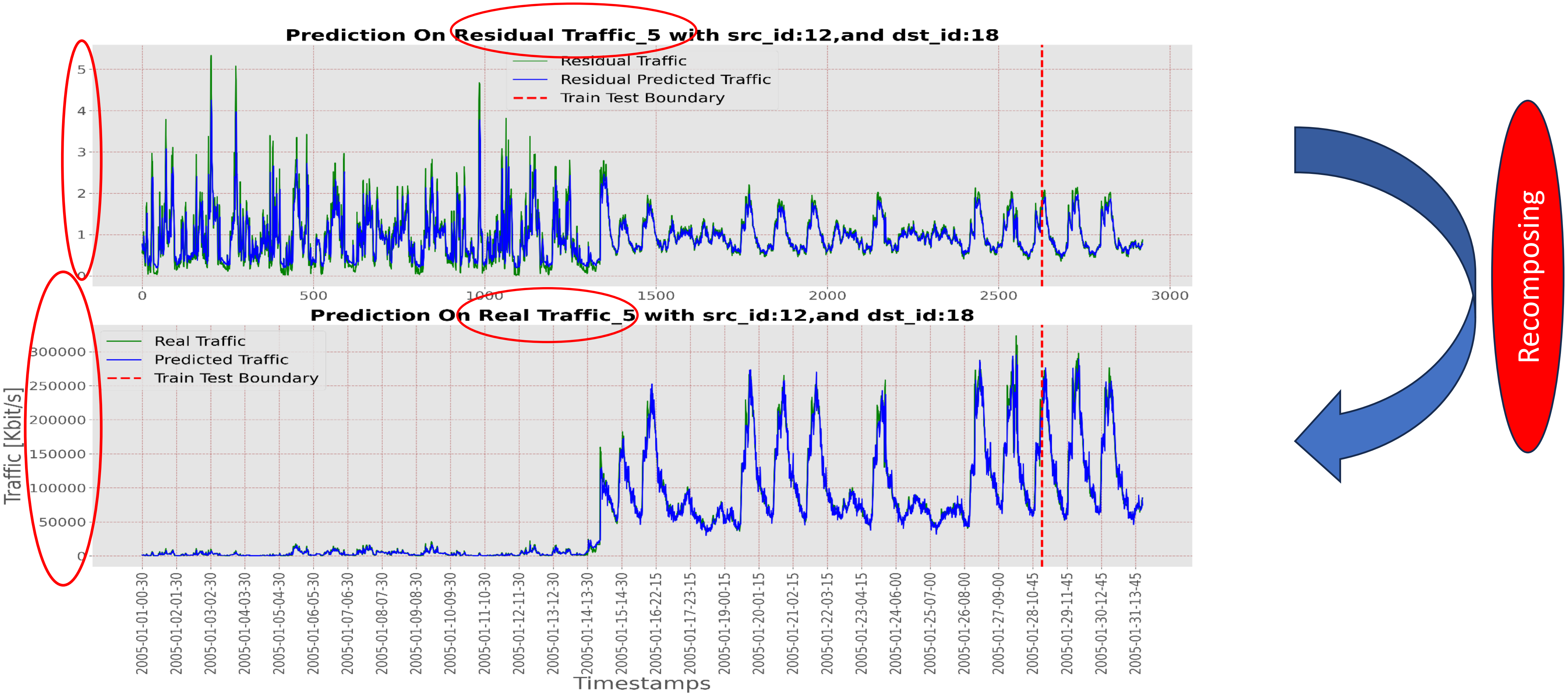
2_clean_traffic_dataframe_without_outliers

Why???

Effect of Outliers

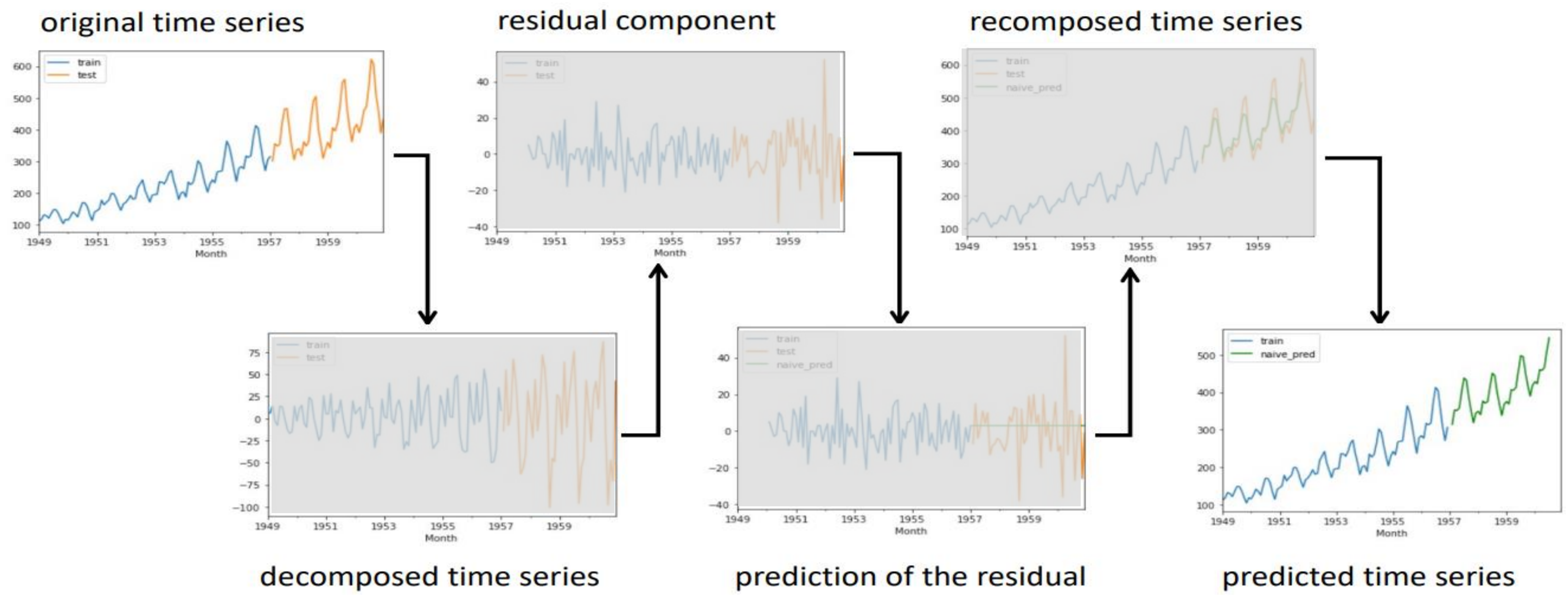
Effect of Normalization

3. Predicting By Residual traffics e f



4. Predicting By Original Traffic

Except decomposing, residual and recomposing section, other steps are similar to predicting by residuals



5. Conclusion

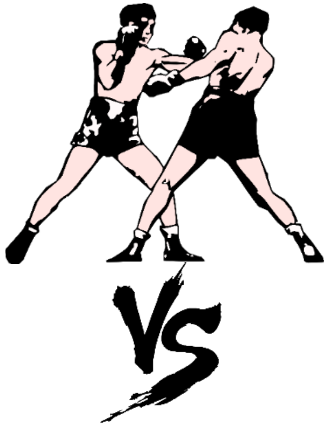
Comparison 1

Effect of Outliers And Normalization

a *Predicting with outliers*

b *Predicting without outlier*

Worse MAE, MSE, MAPE, Training time



Better MAE, MSE, MAPE, Training time

c *Predicting without normalization*

d *Predicting with normalization*

Evidence 1

Effect of outliers

Traffic_id	Src_id	Dst_id	Model	Prediction_type	Input	Training_time[s]	MAE	MSE	MAPE[%]
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	with_outliers	111	0.107067	0.025806	13.890752
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	without_outliers	75	0.103827	0.026435	9.96549412
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	normalized_without_outliers	77	0.017821	0.000823	9.20582459

Prediction on residual traffic by Local CONV_LSTM

Effect of Normalization

Traffic_id	Src_id	Dst_id	Model	Prediction_type	Input	Training_time[s]	MAE	MSE	MAPE[%]
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	with_outliers	111	0.107067	0.025806	13.890752
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	without_outliers	75	0.103827	0.026435	9.96549412
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	normalized_without_outliers	77	0.017821	0.000823	9.20582459

Prediction on residual traffic by Local CONV_LSTM

10. Conclusion

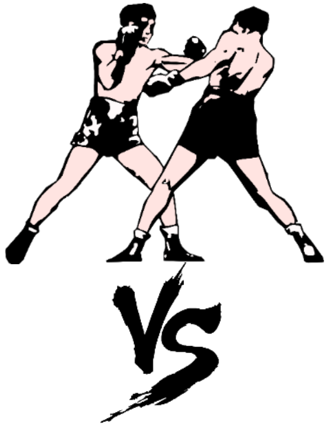
Comparison 2

Effect of residual traffic And Model type

a *Predicting on Original traffic*

Worse MAE, MSE, MAPE, Training time

c *Predicting with RNN*



b *Predicting on Residual traffic*

Better MAE, MSE, MAPE, Training time

d *Predicting with CONV_LSTM*

5. Conclusion

Evidence 2

Prediction on residual traffic by Local CONV_LSTM

Traffic_id	Src_id	Dst_id	Model	Prediction_type	Input	Training_time[s]	MAE	MSE	MAPE[%]
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	with_outliers	111	0.107067	0.025806	13.890752
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	without_outliers	75	0.103827	0.026435	9.96549412
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	normalized_without_outliers	77	0.017821	0.000823	9.20582459

Prediction on original traffic by Local CONV_LSTM

Traffic_id	Src_id	Dst_id	Model	Prediction_type	Input	Training_time[s]	MAE	MSE	MAPE[%]
5	12	18	Model_2: CONV_LSTM	Prediction_on_Original	with_outliers	97.0	129941.104045	2.188556e+10	96.448757
5	12	18	Model_2: CONV_LSTM	Prediction_on_Original	without_outliers	238.0	93226.757085	1.307787e+10	64.668856
5	12	18	Model_2: CONV_LSTM	Prediction_on_Original	normalized_without_outliers	77.0	0.055139	4.539699e-03	14.349007

Effect of residual

Prediction on residual traffic by Local RNN

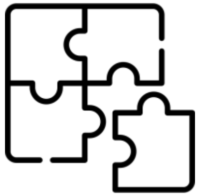
Traffic_id	Src_id	Dst_id	Model	Prediction_type	Input	Training_time[s]	MAE	MSE	MAPE[%]
5	12	18	Model_1: RNN	Prediction_on_residual	with_outliers	91	0.121585	0.026951	15.5907099
5	12	18	Model_1: Rnn	Prediction_on_residual	without_outliers	97	0.101539	0.021906	10.5894383
5	12	18	Model_1: Rnn	Prediction_on_residual	normalized_without_outliers	89	0.086909	0.008859	67.7297785

Prediction on residual traffic by Local CONV_LSTM

Traffic_id	Src_id	Dst_id	Model	Prediction_type	Input	Training_time[s]	MAE	MSE	MAPE[%]
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	with_outliers	111	0.107067	0.025806	13.890752
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	without_outliers	75	0.103827	0.026435	9.96549412
5	12	18	Model_2: CONV_LSTM	Prediction_on_residual	normalized_without_outliers	77	0.017821	0.000823	9.20582459

Effect of Model

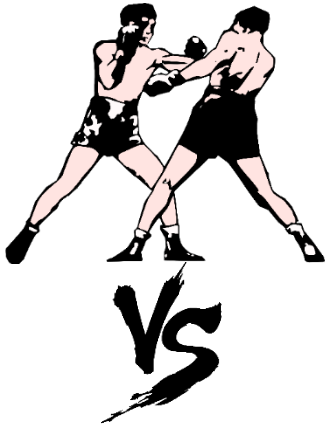
5. Conclusion



Comparison 3
Global vs Local

a *Predicting by Global*

- ***Worse MAE, MSE, MAPE, Training time***
- ***Better Generalization***



b *Predicting by Local*

- ***Better MAE, MSE, MAPE, Training time***
- ***Worse generalization***

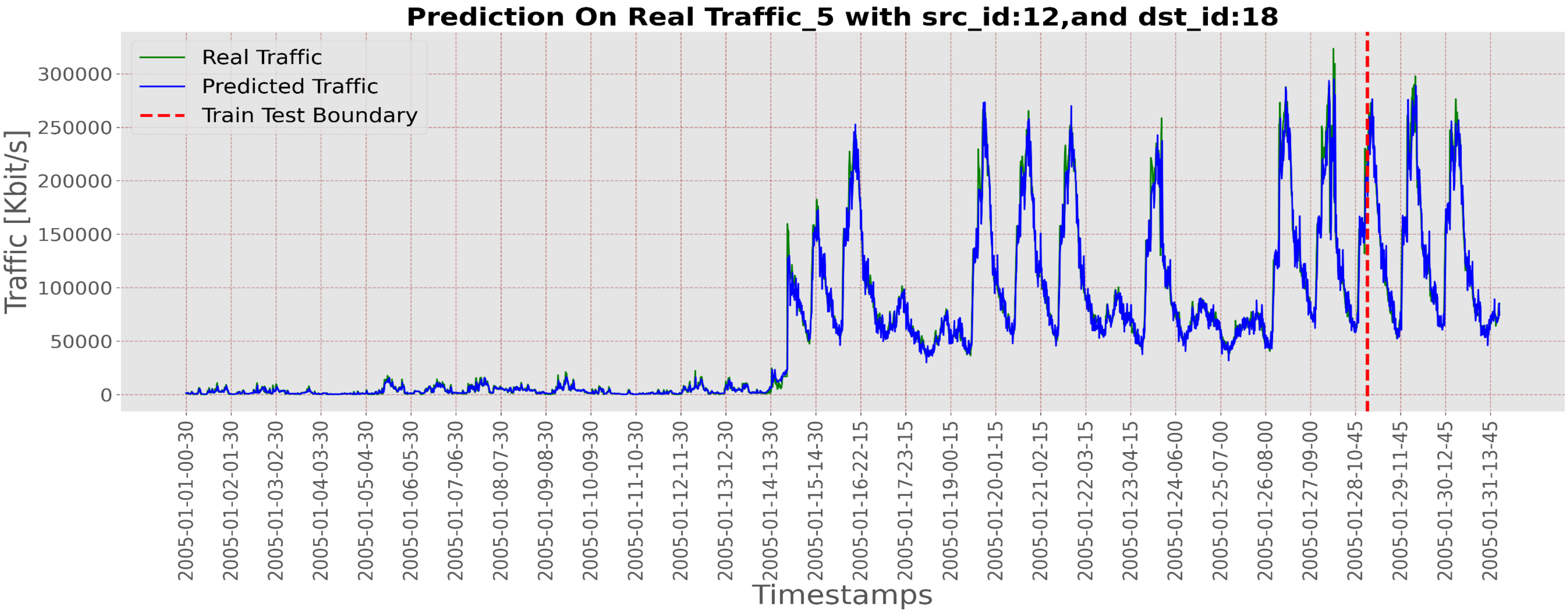
Evidence 3

Traffic_id	Src_id	Dst_id	Model	Prediction_type	Input	Training_time[s]	MAE	MSE	MAPE[%]
5	12	18	Model_2: CONV_LSTM	Prediction_on_Original	normalized_without_outliers	77.0	0.055139	4.539699e	14.349007

Traffic_id	Src_id	Dst_id	Model	Prediction_type	Input	Training_time[s]	MAE	MSE	MAPE[%]
5	12	18	Model_1: Global Transformer	Prediction_on_Original	Without_outliers_normalized	1705	0.718665	1.064684	83.2439744

5. Conclusion

Our Prediction Graph



6. Future Recommendation

1_ Use more advanced anomaly detection models such as decision-based trees_XGBoost

2_Use more advanced models with deeper layers

3_Extend data set and use data augmentation models for more generalization

4_Incorporate Advanced Feature Engineering

THANK YOU

WE APPRECIATE YOUR FEEDBACK