

Prediksi lalu lintas pada weekend di Jalan Raya Pajajaran - Bogor

Faddli Lindra Wibowo



TABLE OF CONTENTS

01

Business Understanding

02

Data Understanding

03

Modeling

04

Evaluation

05

Results

06

Reference





Business Understanding

Problem Statement



Kota Bogor berdasarkan penelitian Global Traffic Scorared 2021 menempati posisi kelima kota termacet di Indonesia



◀ Dampak Kemacetan Lalu Lintas ▶



Kerugian Waktu

Karena kecepatan perjalanan yang rendah



Pemborosan Energi

Karena pada kecepatan rendah konsumsi bahan bakar lebih tinggi



Polusi Udara

Karena pada kecepatan rendah konsumsi energi lebih tinggi



Stres Pengguna Jalan

Karena kecepatan perjalanan yang rendah dapat meningkatkan stress pengguna jalan



Data Source



Dataset yang digunakan pada project ini adalah crowdsourced/user generated data program Waze for Cities dari Waze. Dataset yang digunakan dalam mengerjakan mini project, yaitu aggregate_median_jams kota Bogor. Data tersebut merupakan hasil agregat dari raw data yang dirilis oleh Waze. Data agregat tersebut memiliki rentang waktu 6 Juli 2022 hingga 6 September 2022



Challenge

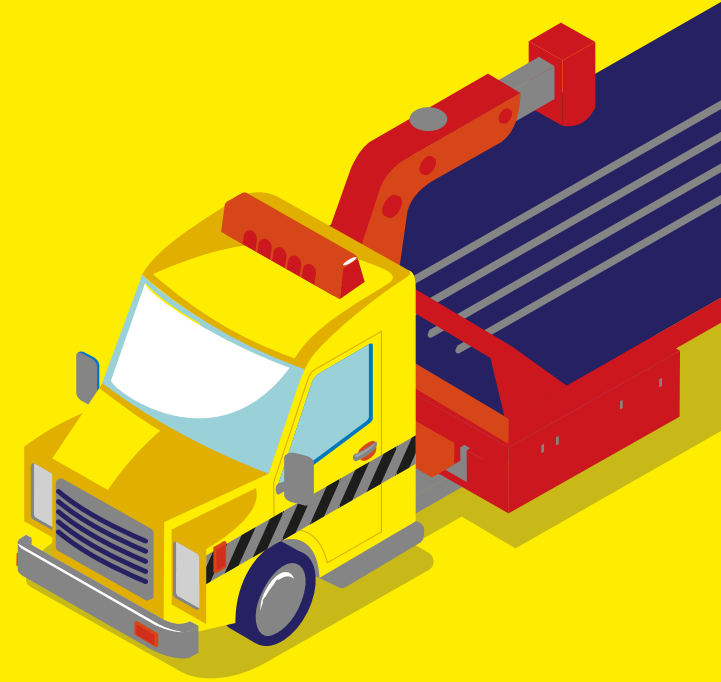
Menentukan algoritma machine learning yang sesuai untuk predictive



Menentukan metrics yang digunakan untuk prediksi



Data berukuran besar



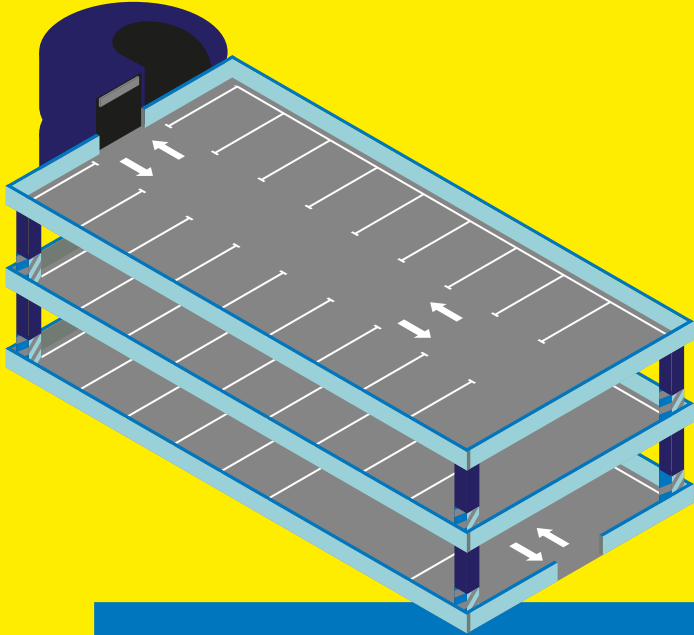
Methodology



Predictive analysis (machine learning model)
untuk memprediksi lalu lintas pada weekend di
Jalan Raya Pajajaran - Bogor



Business Benefits



- Waktu perjalanan lebih efisien
- Menghemat bahan bakar
- Meningkatkan produktivitas
- Meningkatkan kualitas udara
- Mengurangi stres dan risiko kecelakaan
- Meningkatkan kenyamanan
- Meningkatkan mobilitas

Methodology



Predictive analysis (machine learning model)
untuk memprediksi lalu lintas pada weekend di
Jalan Raya Pajajaran - Bogor



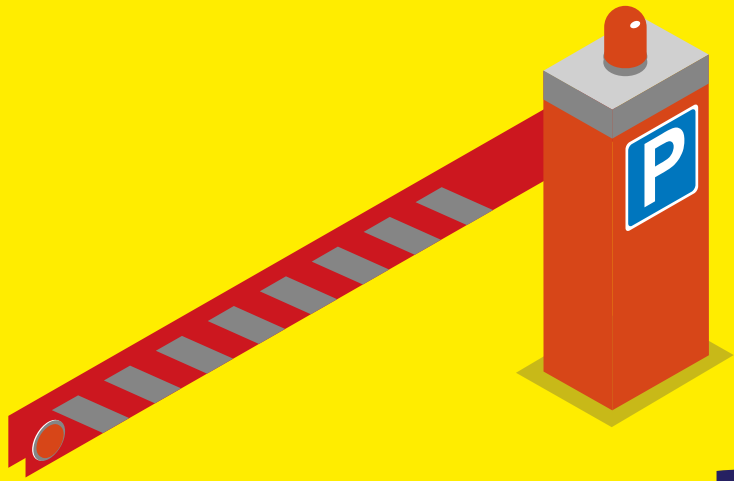
Expected outcomes



Machine learning model

Machine learning model untuk memprediksi lalu lintas pada weekend di Jalan Raya Pajajaran - Bogor





02

Data Understanding

Data Understanding

Dataset yang digunakan dalam mengerjakan mini project, yaitu aggregate_median_jams kota Bogor. Data tersebut merupakan hasil agregat dari raw data yang dirilis oleh Waze

Data agregat tersebut memiliki rentang waktu 6 Juli 2022 hingga 6 September 2022

Data memiliki atribut time, kemendagri kabupaten kode, kemendagri kabupaten nama, street, level, median length, median delay, median speed kmh, total records, id , date, dan geometry

Deskripsi Level

0 = free flow speed to 80% of free flow speed

1 = 80% to 61% of free flow speed

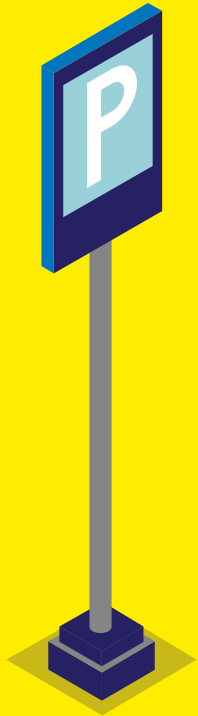
2 = 60% to 41%

3 = 40% to 21%

4 = 20% to 1%

5 = blocked road

◀ Data Cleansing and Preprocessing ▶



Cek Data Null dan
hapus dan Nul



Cek data duplikat
dan hapus



Hapus kolom yang
tidak diperlukan



Data selesai di
bersihkan dan siap
digunakan



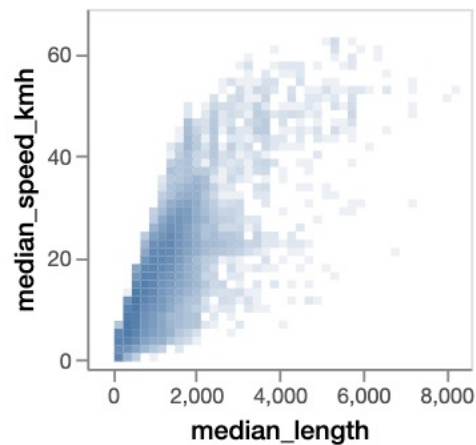
Filter data berdasarkan
nama jalan dan hari
weekend



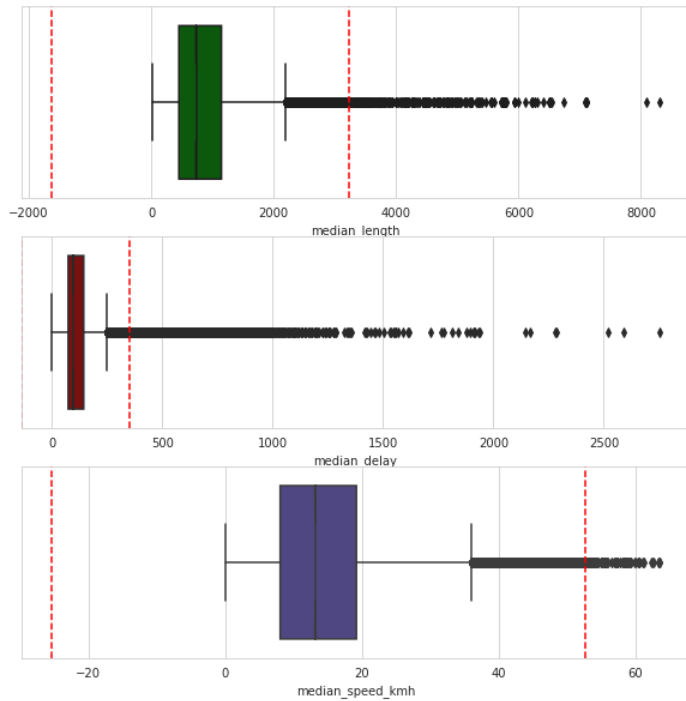
Menentukan weekend
dan weekday

EDA (Exploratory Data Analysis)

Korelasi Heatmap menunjukkan bahwa median_speed_kmh memiliki korelasi sebesar 0.8 dengan median length



EDA (Exploratory Data Analysis)

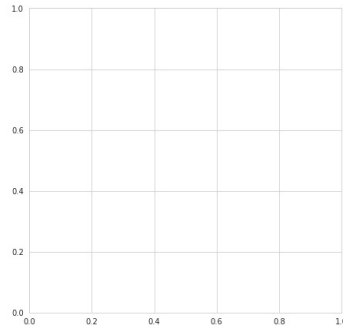
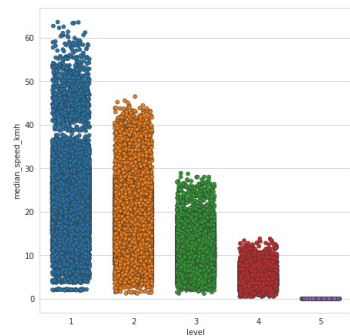
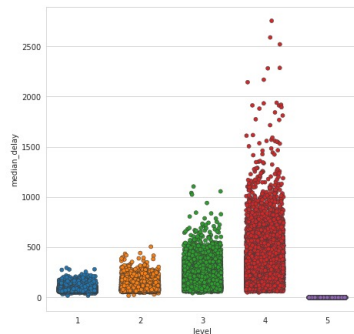
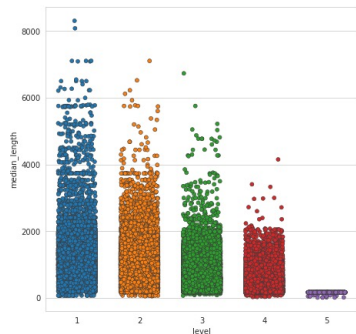


Garfik yang menunjukkan tingkat outlier pada median_delay, median_speed_kmh dan median_length

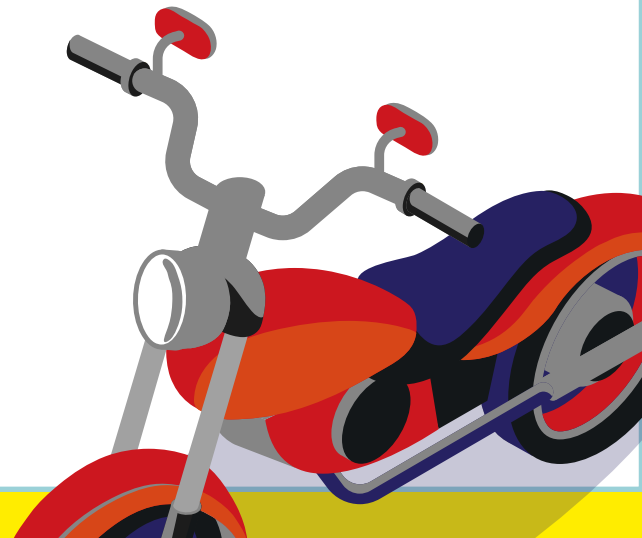


EDA (Exploratory Data Analysis)

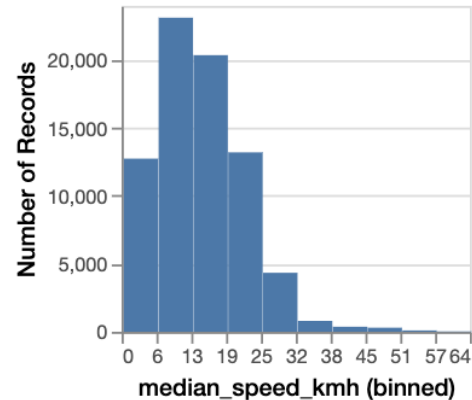
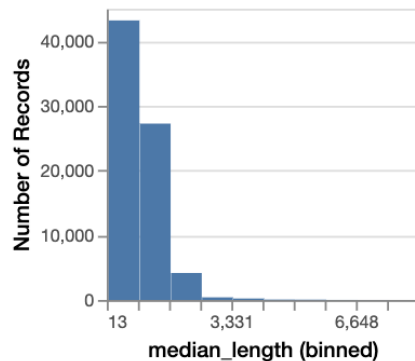
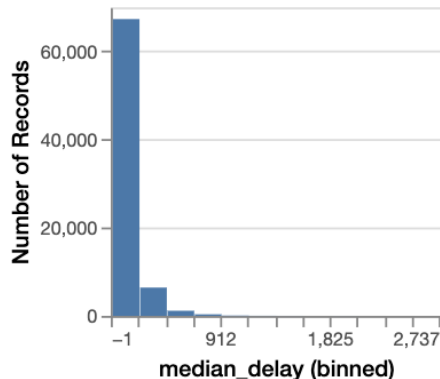
Numerical Attributes vs. Target



Perbandingan level dengan median_delay, median_speed_kmh dan median_length



EDA (Exploratory Data Analysis)

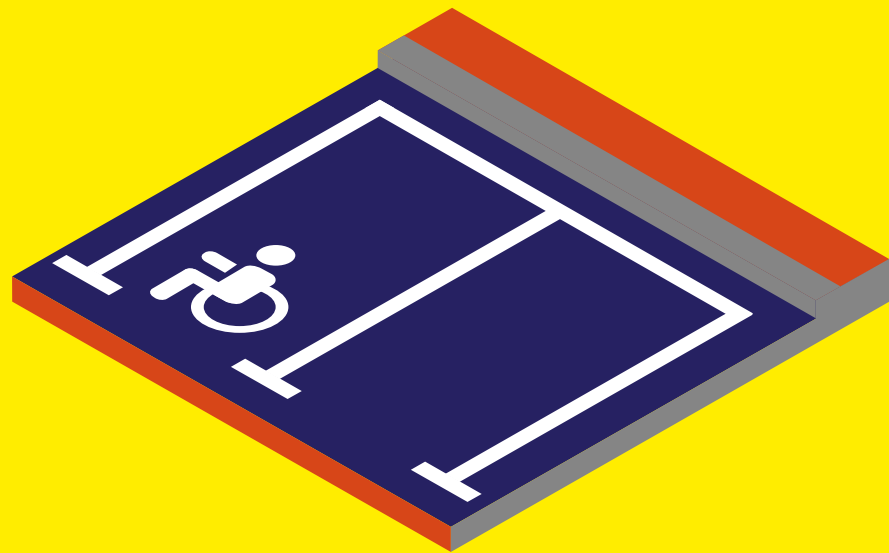


	level	median_length	median_delay	median_speed_kmh
count	100529.000000	100529.000000	100529.000000	100529.000000
mean	2.421013	848.935556	131.950621	14.199692
std	0.999856	550.099405	110.124851	7.975619
min	1.000000	13.000000	-1.000000	0.000000
25%	2.000000	453.000000	75.000000	8.050000
50%	2.000000	735.000000	97.000000	13.190000
75%	3.000000	1147.000000	145.000000	19.210000
max	5.000000	8307.000000	2754.000000	63.614998

Tabel yang menampilkan deskripsi dari data median_delay, median_speed_kmh dan median_length. Mencakup jumlah data, mean, standar deviasi, data minimum dan maximum

03

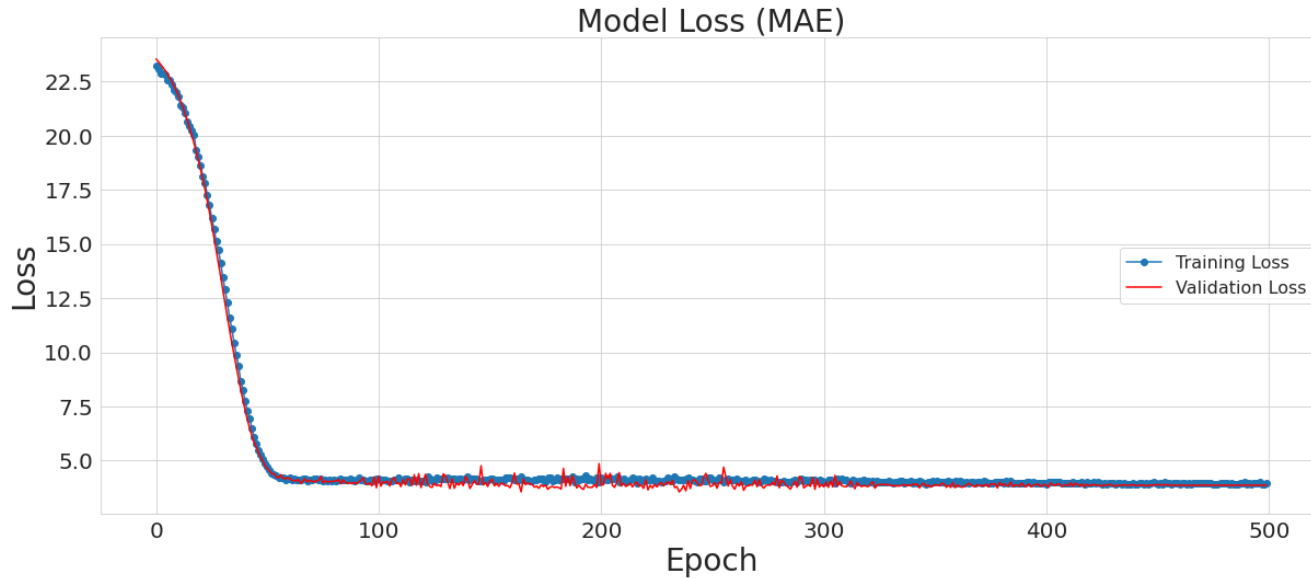
Modeling



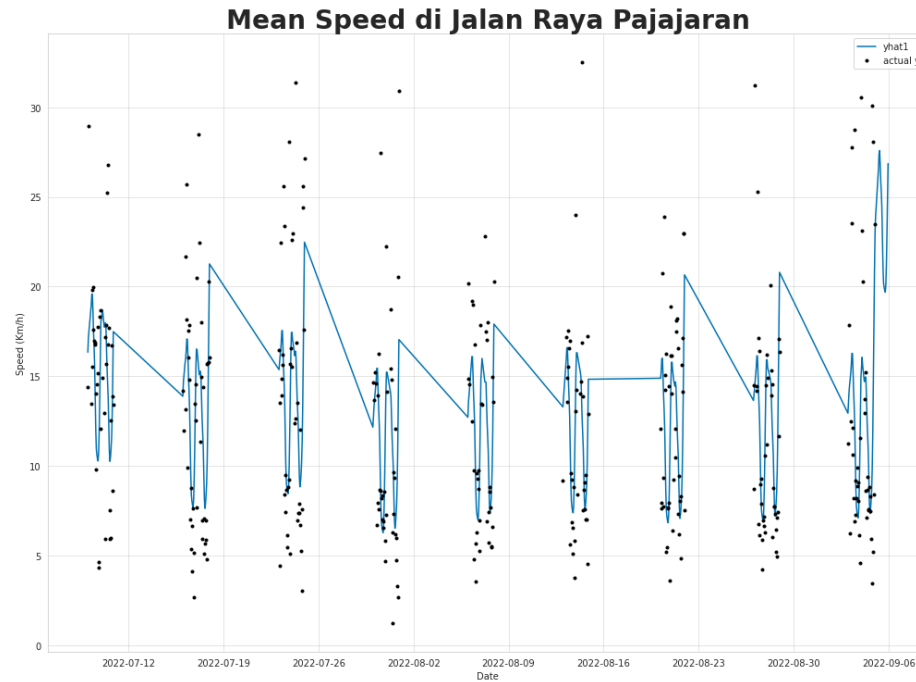
Machine Learning Model

- Filter data yang digunakan adalah nama jalan Raya Pajajaran dan hari weekend
- Pada machine learning modeling menggunakan atribut time (ds) dan median_speed_kmh (y)
- Library menggunakan Neuralprophet
- Frekuensi dataset adalah '1 jam'.
- Epoch yang digunakan 500
- Batch_size = 512
- Jumlah of hidden layer = 8 dan Dimension of hidden layers = 4
- learning_rate=0.01

Training dan Validation Loss



Grafik menunjukkan hasil yang cukup bagus dimana antar train dan validation loss memiliki nilai yang tidak jauh berbeda. Nilai Loss menurun hingga nilai 3.8 dari epoch 50 dan stabil sampai epoch 500



Grafik menunjukkan data aktual dan data prediksi menggunakan neuralprophet.

04

Evaluation



Mean Absolute Error (MAE)

MAE atau Mean Absolute Error menunjukkan nilai kesalahan rata-rata yang error dari nilai sebenarnya dengan nilai prediksi. MAE sendiri secara umum digunakan untuk pengukuran prediksi error pada analisis time series. Hasil MAE dari modeling machine learning prediksi lalu lintas pada weekend sebesar 3,97 dimana hasil ini sudah bagus

Rumus dari MAE sendiri didefinisikan sebagai berikut :

$$MAE = \sum \frac{|Y' - Y|}{n}$$

Y ' = Nilai Prediksi

Y = Nilai Sebenarnya

n = Jumlah Data

Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE), adalah jumlah dari kesalahan kuadrat atau selisih antara nilai sebenarnya dengan nilai prediksi yang telah ditentukan. Hasil RMSE dari modeling machine learning prediksi lalu lintas pada weekend sebesar 5,13 dimana hasil ini sudah bagus

Rumus formula RMSE adalah sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y})^2}{n}}$$

05

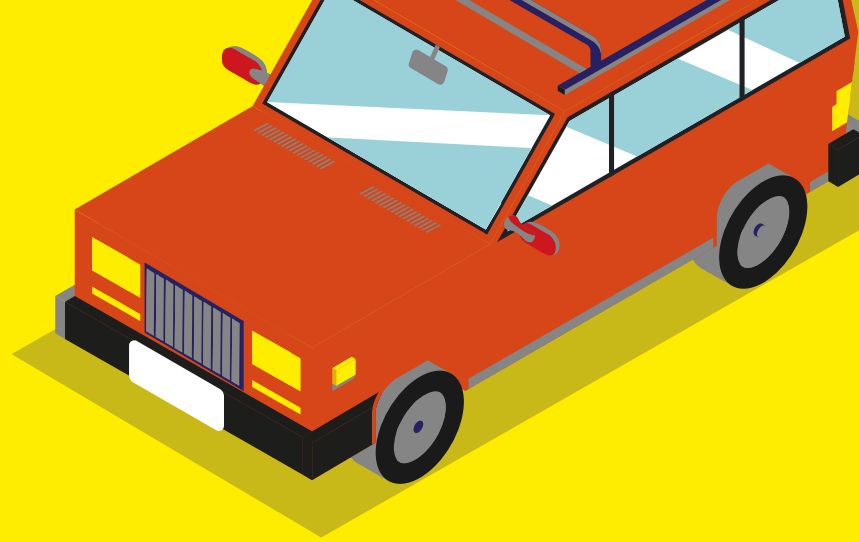
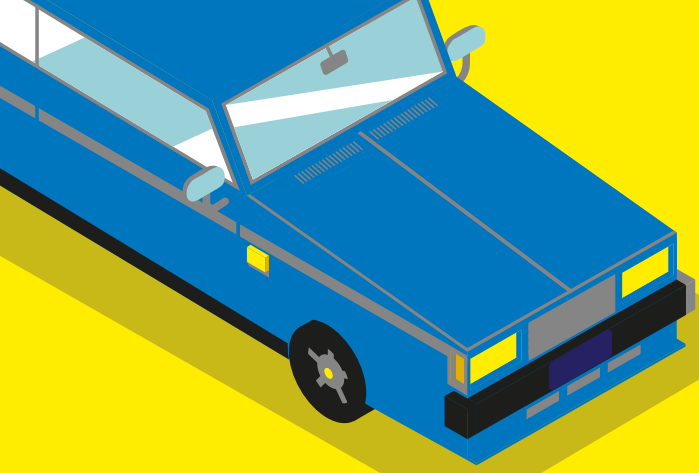
Result



Hasil Prediksi Untuk 1 Hari Ke depan

	ds	yhat1
0	2022-09-05 00:00:00	23.018282
1	2022-09-05 01:00:00	23.956078
2	2022-09-05 02:00:00	24.422817
3	2022-09-05 03:00:00	24.967113
4	2022-09-05 04:00:00	25.481348
5	2022-09-05 05:00:00	25.906979
6	2022-09-05 06:00:00	26.573635
7	2022-09-05 07:00:00	27.387516
8	2022-09-05 08:00:00	27.583624
9	2022-09-05 09:00:00	26.860676
10	2022-09-05 10:00:00	25.920092
11	2022-09-05 11:00:00	25.279366
12	2022-09-05 12:00:00	24.424057
13	2022-09-05 13:00:00	22.854473
14	2022-09-05 14:00:00	21.157053
15	2022-09-05 15:00:00	20.205807
16	2022-09-05 16:00:00	19.948553
17	2022-09-05 17:00:00	19.794250
18	2022-09-05 18:00:00	19.685844
19	2022-09-05 19:00:00	20.056757
20	2022-09-05 20:00:00	21.134430
21	2022-09-05 21:00:00	22.845762
22	2022-09-05 22:00:00	24.959440
23	2022-09-05 23:00:00	26.862450

Tabel menunjukkan prediksi untuk 1 hari kedepan yaitu pada tanggal 05-09-2022 (Senin). Pada tanggal tersebut dimulai pada jam 00.00 sampai jam 23.00. Hasil prediksi yaitu kolom yhat1 menunjukkan speed dalam satuan km/h. Kecepatan tidak ada menunjukkan dibawah angka 10 yang berarti ada kemacetan



06

Reference

Reference

1. <https://news.detik.com/berita-jawa-barat/d-5894932/bogor-jadi-kota-termacet-kelima-di-indonesia-begini-respons-bima-arya>. Download Apps Detikcom Sekarang
<https://apps.detik.com/detik/>
2. <https://id.wikipedia.org/wiki/Kemacetan>
3. <https://neuralprophet.com/contents.html>
4. <https://www.kaggle.com/code/manovirat/timeseries-using-prophet-hyperparameter-tuning>
5. <https://www.amazon.com/Python-Data-Cleaning-Cookbook-techniques/dp/1800565666>





Terima Kasih ▶

DO YOU HAVE ANY QUESTIONS?

bfaddly@gmail.com

<https://github.com/FaddliLWibowo?tab=repositories>

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**