



Studi Kasus Seleksi DSLS 2023

Disajikan oleh: Michael Yulius Sitanggang

 michaelsitanggang18@gmail.com

Latar Belakang

Mayoritas warga New York bergantung kepada layanan transportasi publik atau layanan taksi. Hanya 22% dari penduduk Manhattan memiliki mobil jika dibandingkan dengan rata-rata 91% rumah tangga di seluruh Amerika Serikat memiliki setidaknya satu mobil. Layanan taksi di New York merupakan jaringan transportasi terbesar keempat di Amerika Serikat.

Sistem tersebut diatur oleh New York City Taxi and Limousine Commission (TLC), sistem yang ada sekarang tidak melayani daerah (boroughs) secara merata. Karena hal tersebut, penduduk lebih memilih layanan taksi ilegal yang memiliki akses dan tarif yang tidak konsisten.

Kata Kunci Question : layanan yellow taxi hanya berpusat di boroughs tertentu atau kurang tersebar secara merata di seluruh kota New York.



Penyebab

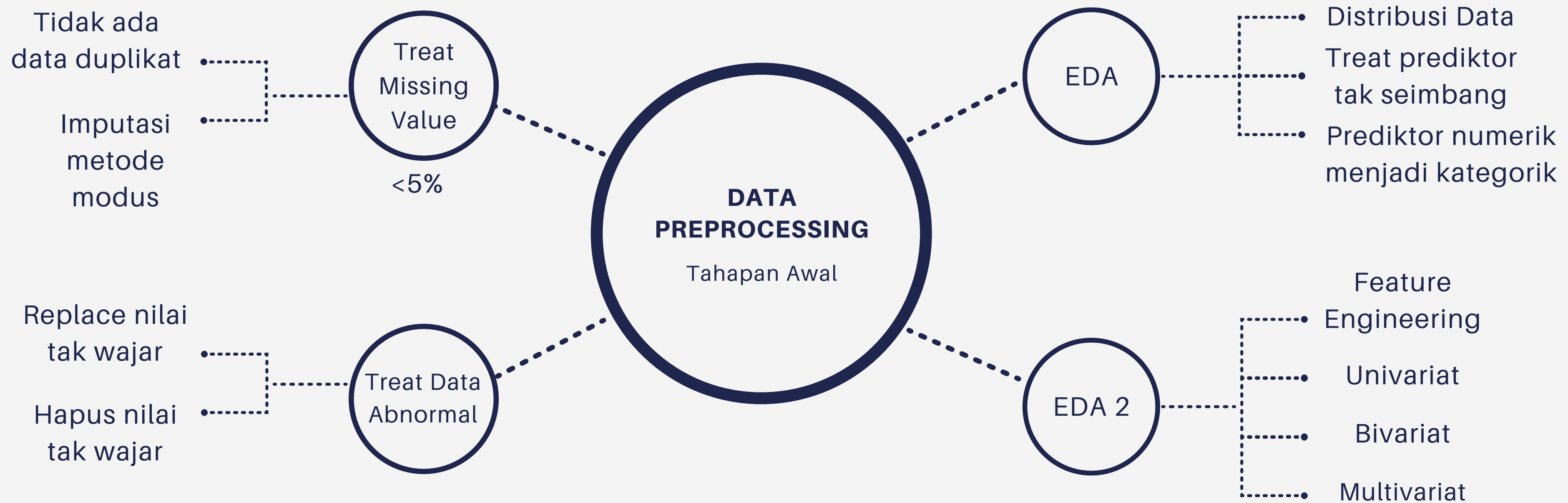
ketidakseimbangan permintaan dan penawaran pada boroughs lain

Rencana 6 Bulan Kedepan

TLC ingin agar layanan yellow taxi mereka tersebar secara merata di seluruh kota New York sehingga dapat menjawab demand yang ada

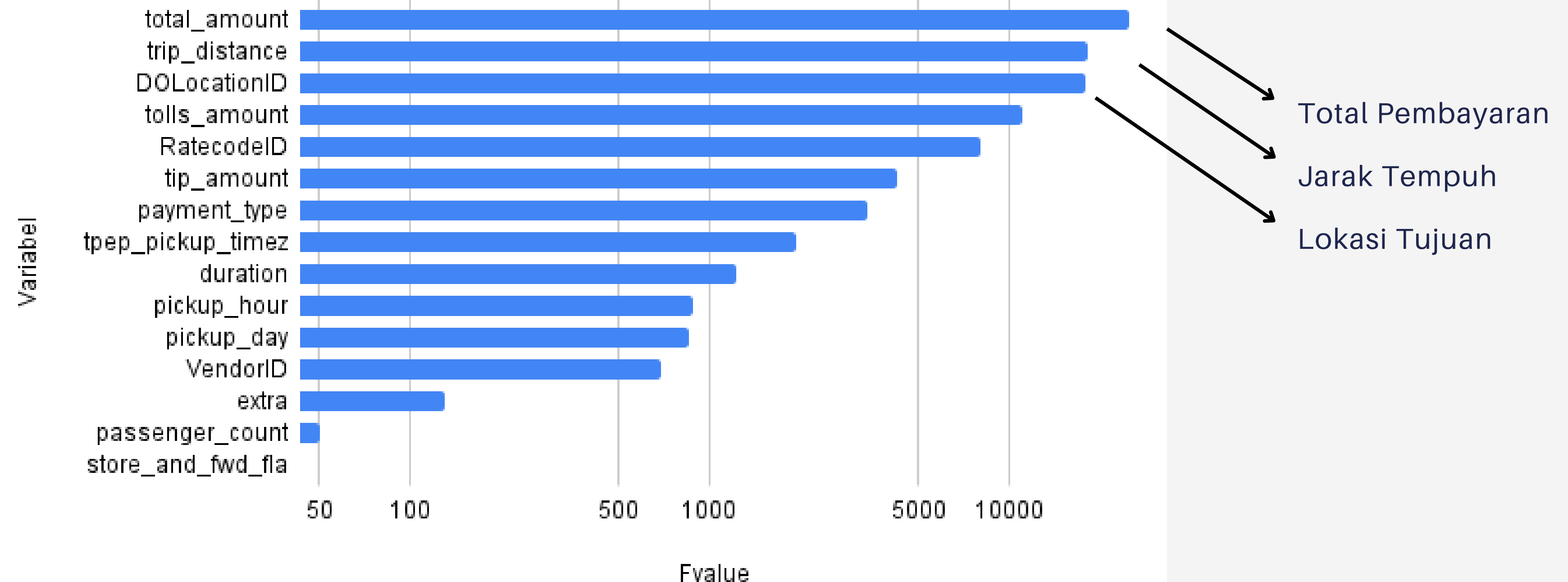
Model Machine Learning

rekomendasi untuk menyelesaikan permasalahan yang ada di TLC

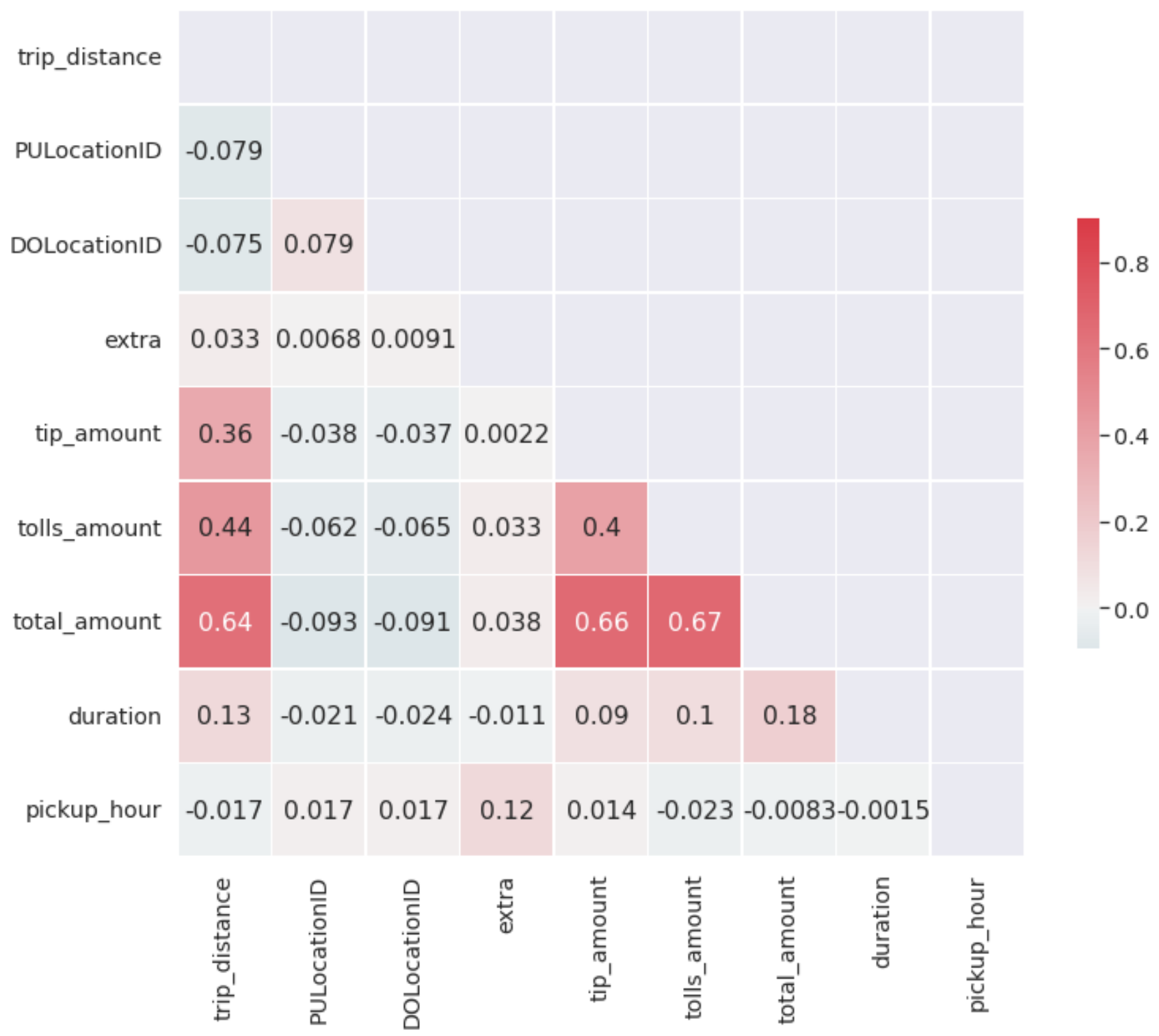


Apa penyebab layanan yellow taxi hanya berpusat di boroughs tertentu sehingga menyebabkan ketidakseimbangan permintaan dan penawaran pada boroughs lain?

Variable Important by Fvalue



Cara yang harus TLC lakukan secara *short-term* (6 bulan) agar layanan yellow taxi tersebar secara merata di seluruh kota New York ?



1

JARAK TEMPUH

Dari plot Korelasi dan secara perhitungan regresi total pembayaran didapatkan bahwa jarak tempuh menjadi faktor penting persebaran TLC seluruh kota NY

SOLUSI = UPGRADE FISIK TAXI UNTUK KETEPATAN WAKTU DROPOFF PELANGGAN DAN KENYAMANAN JIKA JARAK TEMPUH TUJUAN CUKUP JAUH

2

LOKASI TUJUAN

Dari plot Korelasi juga ditunjukkan total tip mempengaruhi total pembayaran. Variabel yang sama juga muncul menjadi faktor pembobot model prediksi total pembayaran

SOLUSI = PELAYANAN DRIVER JUGA MENJADI BAGIAN PERUBAHAN YAITU MEMBERI PELAYANAN PRIMA SAAT MENERIMA TIP / TIDAK MENERIMA TIP.

3

BIAYA TAMBAHAN

Dari plot Korelasi juga ditunjukkan tambahan biaya tol mempengaruhi total pembayaran. Variabel yang sama juga muncul menjadi faktor pembobot model prediksi total pembayaran

SOLUSI = PENINJAUAN KEMBALI SETIAP BIAYA TAMBAHAN YANG AKAN DIBAYAR PELANGGAN, INI JUGA SALAH SATU CARA BERSAING TERHADAP LAYANAN TAKSI ILEGAL DI NEW YORK.

Rekomendasi



Rekomendasi model Machine Learning untuk memprediksi Total Pembayaran adalah ElasticNet (Linear)

Variabel yang mempengaruhi dalam pembobotan model antara lain

- Jarak Tempuh
- Biaya Tambahan Akses Tol
- Tambahan Tip

Tingkat Error Train Model sebesar 2.901203 dan Test Model sebesar 2.907404 , tidak terjadi overfitting dalam model

Analisis Error menggambarkan bahwa terjadi heteroskedastisitas dan error tidak berdistribusi normal maka perlu dikembangkan lanjut menggunakan model non linear

TERIMA KASIH



DRIVE DATA

https://drive.google.com/drive/folders/1LOadj_Z3FJVkvFdUnqNbmXecB8GVaOzu?usp=share_link

SCRIPT .ipynb

https://colab.research.google.com/drive/1n1NylsQO75q7BTX6PTruhqD1wN6PI-uN?usp=share_link