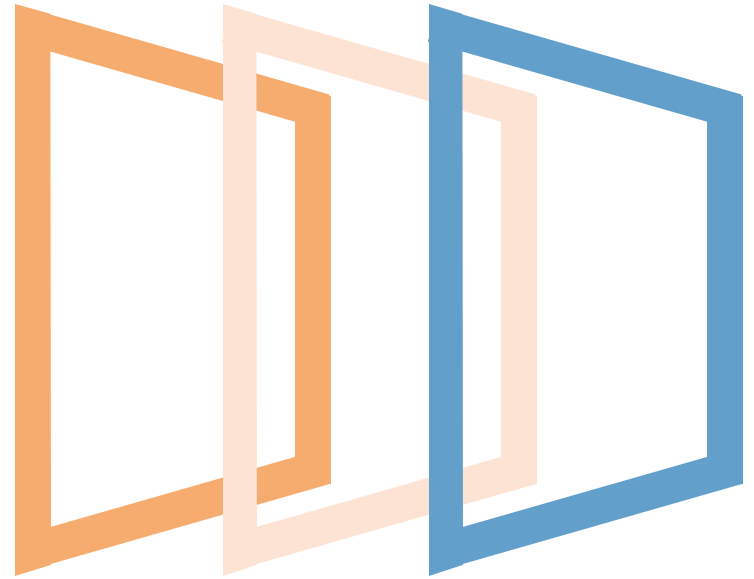


Taxi Chicago

Equipo 3

Octubre 6

minsoit



An Indra company

Índice

- 01. Consideraciones
- 02. Modelos considerando 3 ventanas
- 03. Modelos considerando 6 ventanas
- 04. Conlcusiones

Consideraciones

01

- Se tomo como umbral aquellos taxis que trabajaron al menos 28 semanas hasta el 31 de julio de 2023.

Al generar los modelos se tuvieron dos escenarios:

- Un sobreajuste, en modelos de ensamble.
- Casos de subajuste, en modelos de regresión lineal y otros.

.

Para mitigar el sobreajuste de los modelos, se tomó en cuenta:

- Creación de nuevas variables
- Correlaciones
- Varianza Inflation Factor (VIF)
- Ventanas de diferente longitud.

02.Modelos: 3 ventanas

02

VIF

Texto

Features	VIF
fare_mean_3	32.065769
km_mean_3	10.752988
sec_mean_3	24.899861
extras_mean_3	4.763239
tips_mean_3	9.174890

minsoit

An Indra company 12

Texto new ideas principal

Features	VIF
fare_mean_3	6.008255
extras_mean_3	3.277440
Ptips_mean_3	4.452123
Ptolls_mean_3	1.015670

minsoit

An Indra company 10

Texto

Features	VIF
FareMiles_mean_3	9.607548
fareSec_mean_3	11.769249
extras_mean_3	3.294004
Ptips_mean_3	4.555079
Ptolls_mean_3	1.018696

minsoit

An Indra company 13

Considerando las siguientes variables de entrada:

'fare_mean_3', 'km_mean_3', 'sec_mean_3', 'extras_mean_3',
'tips_mean_3'

	R2 trainig	R2 validation
Regresión Lineal	▪ 0.62	▪ 0.62
Random Forest	▪ 0.94	▪ 0.61
Red Neuronal 1	▪ 0.61	▪ 0.64
Red Neuronal 2	▪ 0.61	▪ 0.64

Considerando las siguientes variables de entrada:

'fare_mean_3', 'extras_mean_3', 'Ptips_mean_3',
'Ptolls_mean_3'

	R2 trainig	R2 validation
Regresión Lineal	▪ 0.61	▪ 0.61
Random Forest	▪ 0.94	▪ 0.59
Red Neuronal 1	▪ 0.61	▪ 0.64
Red Neuronal 2	▪ 0.61	▪ 0.64

Considerando las siguientes variables de entrada:

'FareMiles_mean_3', 'fareSec_mean_3', 'extras_mean_3',
'Ptips_mean_3', 'Ptolls_mean_3'

	R2 trainig	R2 validation
Regresión Lineal	▪ 0.52	▪ 0.52
Random Forest	▪ 0.94	▪ 0.57
Red Neuronal 1	▪ 0.70	▪ 0.74
Red Neuronal 2	▪ 0.74	▪ 0.78

02.Modelos: 6 ventanas

03

VIF

Text

Features	VIF
fare_mean_6	33.945678
km_mean_6	10.530175
sec_mean_6	26.883448
extras_mean_6	6.342012
tips_mean_6	10.521008

minsoit

An Indra company

Texto novo ideia principal

Features	VIF
fare_mean_6	6.851988
extras_mean_6	3.953275
Ptips_mean_6	4.691093
Ptolls_mean_6	1.023133

minsoit

An Indra company

Text

Features	VIF
FareMiles_mean_6	9.731794
fareSec_mean_6	12.008068
extras_mean_6	3.979643
Ptips_mean_6	4.844745
Ptolls_mean_6	1.026564

minsoit

An Indra company

Considerando las siguientes variables de entrada:

`['fare_mean_6', 'km_mean_6', 'sec_mean_6', 'extras_mean_6', 'tips_mean_6']`

	R2 trainig	R2 validation
Regresión Lineal	▪ 0.62	▪ 0.61
Random Forest	▪ 0.94	▪ 0.60
Red Neuronal 1	▪ 0.63	▪ 0.67
Red Neuronal 2	▪ 0.63	▪ 0.67

Considerando las siguientes variables de entrada:

'fare_mean_6', 'extras_mean_6', 'Ptips_mean_6', 'Ptolls_mean_6'

	R2 trainig	R2 validation
Regresión Lineal	▪ 0.62	▪ 0.61
Random Forest	▪ 0.94	▪ 0.59
Red Neuronal 1	▪ 0.63	▪ 0.67
Red Neuronal 2	▪ 0.63	▪ 0.67

Considerando las siguientes variables de entrada:

'FareMiles_mean_6',

'fareSec_mean_6','extras_mean_6', 'Ptips_mean_6', 'Ptolls_mean_6'

R2 trainig

R2 validation

Regresión Lineal

▪ 0.52

▪ 0.52

Random Forest

▪ 0.94

▪ 0.57

Red Neuronal 1

▪ 0.74

▪ 0.79

Red Neuronal 2

▪ 0.74

▪ 0.77

Selección de modelos

04

Los mejores modelos fueron generados por estructuras d redes neuronales.

Para el modelo de ventanas de tiempo de 3 semanas, se tomaron las siguientes configuraciones.

- Dos caspas de neuronas (30 y 10)
- optimizer='rmsprop'
- loss='mse'
- early_stopping con patience=5
- activation function: relu

Para el modelo de ventanas de tiempo de 6 semanas, se tomaron las siguientes configuraciones.

- Una capa de 64 neuronas
- optimizer='rmsprop'
- loss='mse'
- early_stopping con patience=5
- activation function: relu

Aplicaciones del modelo

1

Estimación de Costos para Usuarios:

- Esto les permite tomar decisiones informadas sobre cómo viajar y planificar su presupuesto.

2

Optimización de Rutas:

- Ayudar a los conductores a seleccionar rutas óptimas y minimizar los costos para los clientes. Esto puede incluir consideraciones como el tráfico, las tarifas de peaje y las tarifas variables.

3

Evaluación de Políticas de Transporte

- Evaluar el impacto de las políticas de transporte en los costos de viaje en taxi y en la movilidad urbana en general.

¡Gracias!

Presentación:

Karen Cruz Garduño

kcruzg@minsait.com

minsait

An Indra company

minsait

Mark Making the way forward

An Indra company