

Elige un viaje o desliza hacia arriba para conocer más



UberX 4

\$11.874

00:34 · 5 min de distancia

⚡ Más rápido



Comfort

\$14.111

00:34



UberXL

\$19.802

00:38

Tiempo de espera

25-40 min

Básica

Programar

Tus artículos

[Ver menú](#)



Llega gratis mañana

Solo en Santiago de Chile

Comprando dentro de las próximas 23 h 42 min

[Ver más formas de entrega](#)

Envío rápido de China a Chile con AliExpress Standard Shipping
Fecha estimada de entrega el 07 ENE

Comprar

Añadir a la cesta



Envío
normal

Se enviará desde nuestra bodega entre el
miércoles 04 de enero y el **martes 10 de enero**.
Lo recibirás en cualquier lugar de Chile entre 1 y 3
días hábiles luego del envío.

Opciones de envío:

- ✓ **Envío en fecha programada** - \$3.990
Envío **gratis** en productos pequeños por compras desde \$29.990.

La disponibilidad puede variar durante la compra.

Costo despacho \$3.990
Desde Mar 13/12 - 09:30 hrs



Retira desde el **Jueves 15 de Dic.**

En Homecenter Puente Alto

[Cambiar](#)

[¿Retira otra persona?](#)

Gratis



Llega el **jueves 15 de dic.** de 8 a 20 h.

[Cambiar fecha](#)

\$ 4.990

11 dic.

@FalabellaAyuda, del 29\11 debía llegar una compra que realice por internet. A la fecha de hoy aún no tengo respuesta por vuestra requerimiento 😡😡😡



2



Mostrar más respuestas

Llevo 1 semana esperando mi cama. Como es posible NADIE me contacte? Necesito anular la compra y no me dejan @SERNAC @FalabellaAyuda chantas que venden humo y venden productos que no tienen. @RipleyChile tienen camas drimkip #falabellaesta

compré en Ikea, llevo pagada 2 cuotas de
llegó y todavía no hacen la anulación de
septiembre 😡😡

Responder Ver traducción

van a enviar la mesa del juego de
arven de nada las sillas sin la mesa... 🙄
a Sernac y si no hay respuesta
da demanda... La peor experiencia que he
... Nivel de incompetencia insuperable
😡😡😡😡

5 h 8 Me gusta Responder Ver traducción

estas (5)

/ Compra efectuado en Agosto
nueve completo, pedí retiro de este de mi
io de ciudad los personajes que atienden, se
taban más preocupados de la hora de almuerzo
ar mi reclamo 3 personas distintas me
lismo. Día, las cuales. Decían que no existía
no trabaja su área de call Center?? Se manda
n jefe Al menos que vigilé que hacen y que
responden a la gente, porque da vergüenza las
respuestas que el cliente devuelva por sus propios medios un

La Promesa:

Entregas confiables



1

Contexto

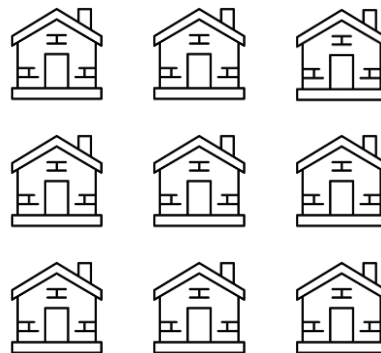


La promesa

Supermercados



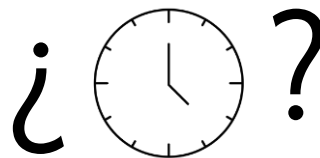
Clientes





La promesa

Promesa



Supermercados

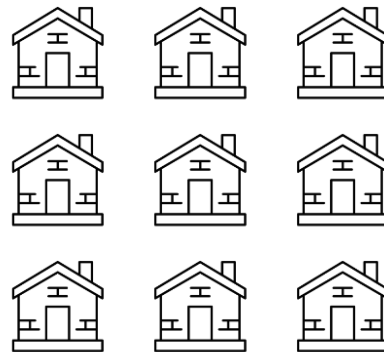


Picker



Pedidos

Clientes

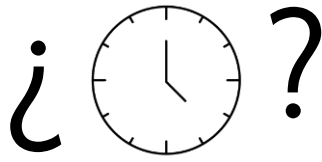




La promesa



Clientes



Promesa

● Planificación

● Confianza

● Fidelización



Objetivo general

Lograr la **mejora de un modelo existente**, con objeto de **diferenciarse** de la competencia y **fidelizar** al cliente final.



Conociendo al equipo



Jorge Hoffmann

Líder proyecto



Mauricio Gómez

Ingeniero de Datos



Constanza Gajardo

Aseguramiento de calidad



Cristian Pérez

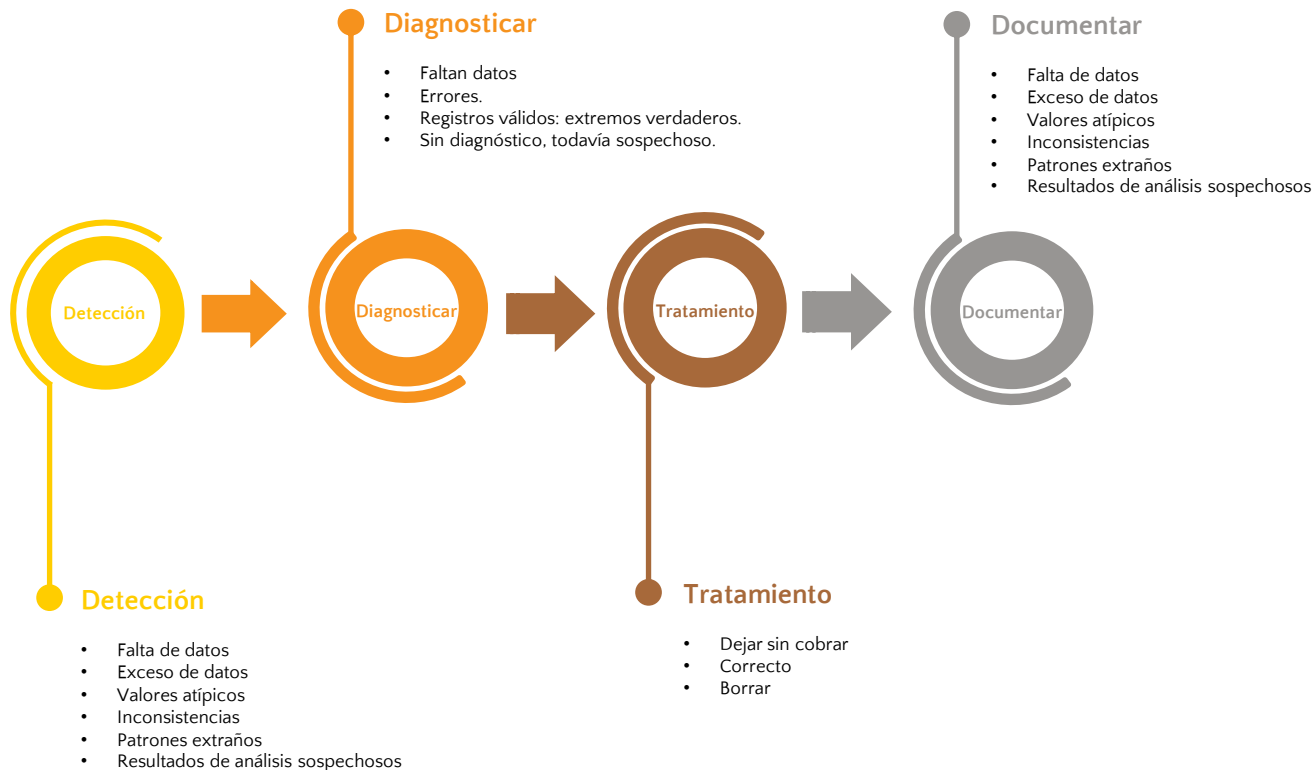
Ingeniero de modelación

2

Manejo de datos



Fases de la limpieza de datos



Data cleaning

Ordenar el set de datos



- Armonizar las fuentes de datos.
- Texto y números alineados (izquierda y derecha respectivamente)
- Cada variable corresponde a una columna.
- Cada observación es una fila.
- No hay filas en blanco.
- Los encabezados de columna son claros y visualmente distintos.
- Se eliminan espacios iniciales.

order_id	lat	lng	dow
6abd5e28d01b7c733fe37485294235dc	1.822.114.109.606.680	-10.031.391.366.664.600	2
0d37a1eea001f01b40232ffd496ff663	18.177.712.202.580.300	-10.035.705.073.096.900	2
528428e9b9eb5d3bd0dbdd8da692b087	1.826.683.205.404.410	-10.031.755.280.890.600	2
d3f7be6d64eae7405b7bf024e04807a2	18.248.884.004.419.400	-1.003.226.215.187.040	2
e0ad6533b6cd9a382ec6c1d473c5096c	1.822.911.027.242.010	-10.032.267.241.783.800	2

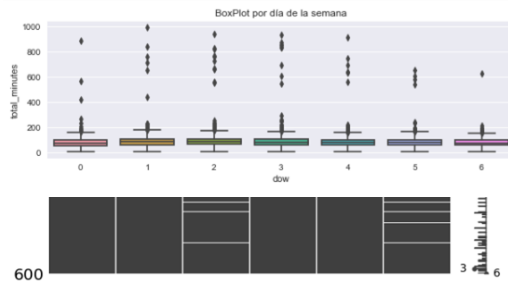
messy	snake case	kebab case	camel case	pascal case	constant case
_Customer_name	customer_name	customer-name	customerName	CustomerName	CUSTOMER_NAME
customer_gender	customer_gender	customer-gender	customerGender	CustomerGender	CUSTOMER_GENDER
customer age	customer_age	customer-age	customerAge	CustomerAge	CUSTOMER_AGE
billingAddress	billing_address	billing-address	billingAddress	BillingAddress	BILLING_ADDRESS
delivery** address	delivery_address	delivery-address	deliveryAddress	DeliveryAddress	DELIVERY_ADDRESS
Item(Description)	item_description	item-description	itemDescription	ItemDescription	ITEM_DESCRIPTION
quantity/Volume	quantity_volume	quantity-volume	quantityVolume	QuantityVolume	QUANTITY_VOLUME
total.price	total_price	total-price	totalPrice	TotalPrice	TOTAL_PRICE



Detección de errores



- Revisión de formatos según tipo de dato y variable.
- ¿Faltan datos?
- ¿Sobran datos?
- Valores atípicos o inconsistentes.



Métodos de detección



- Estadísticas de resumen.
- Exploración gráfica de distribuciones de frecuencia.
- Comprobación de cuestionamientos mediante algoritmos.

- ¿los atributos de tipo de coordenadas tienen valores positivos y negativos?
R: Debemos manipular un poco los valores de los atributos a analizar, para eso se procedera a eliminar los . y convertir a int64

```
store['lat'] = store['lat'].map(lambda x: x.replace('.', '').astype('int64'))  
store['lng'] = store['lng'].map(lambda x: x.replace('.', '').astype('int64'))
```

```
print("""De un total de {} registros realizamos una inspeccion  
para revisar la latitud, se detecta que {} registros son  
mayores o igual a 0, mientras que {} registros son menores a 0""")  
.format(store.shape[0], store[store['lat'] >= 0]  
.shape[0], store[store['lat'] < 0].shape[0]))
```

De un total de 113 registros realizamos una inspeccion para revisar la latitud, se detecta que 113 registros son mayores o igual a 0, mientras que 0 registros son menores a 0

```
print("""De un total de {} registros realizamos una inspeccion  
para revisar la longitud, se detecta que {} registros son  
mayores o igual a 0, mientras que {} registros son menores a 0""")  
.format(store.shape[0], store[store['lng'] >= 0]  
.shape[0], store[store['lng'] < 0].shape[0]))
```

De un total de 113 registros realizamos una inspeccion para revisar la longitud, se detecta que 0 registros son mayores o igual a 0, mientras que 113 registros son menores a 0

Tratamiento de datos.

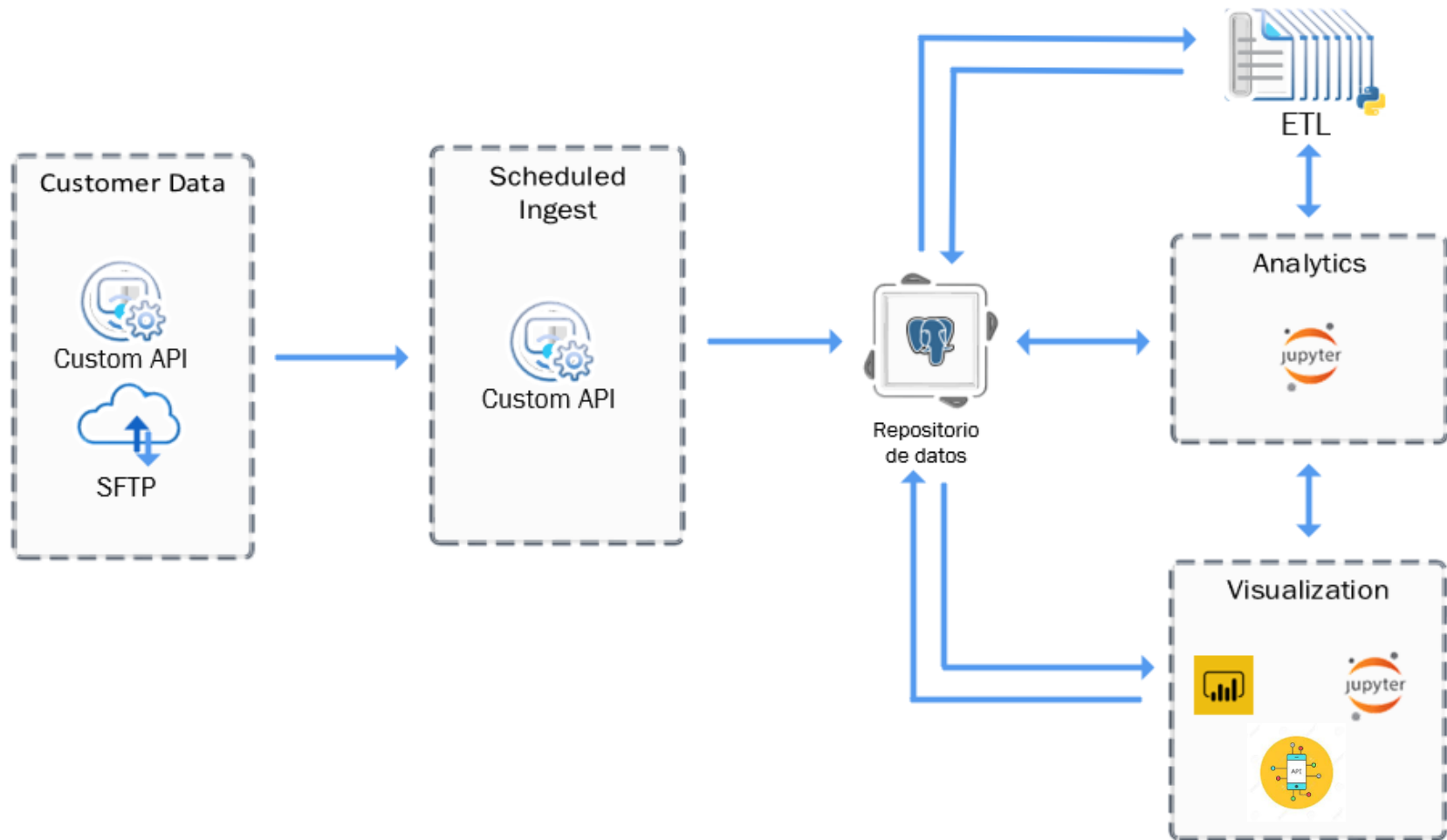
- Transformación
- Imputación.
- Eliminar.
- Eliminar duplicados.
- No modificar.

	shopper_id	seniority	found_rate	picking_speed	accepted_rate	rating
total_missing	0.0	0.0	36.0	0.0	6.0	34.000
percent_missing	0.0	0.0	6.0	0.0	1.0	5.667

	shopper_id	seniority	found_rate	picking_speed	accepted_rate	rating
total_missing	0.0	0.0	0.0	0.0	0.0	0.0
percent_missing	0.0	0.0	0.0	0.0	0.0	0.0

	total_minutes
total_missing	0.0
percent_missing	0.0





3

Modelos

Error del modelo actual

El error actual se medirá por dos ítems:

- ☉ **Mean Squared Error:** promedio de los errores al cuadrado. Particularmente útil ya que castiga más las observaciones más alejadas. Sensible a outliers.
- ☉ **Median absolute error:** Mediana de los errores absolutos. Es insensible a outliers.

Para el modelo actual tenemos:

El error cuadrático medio es de **1.926** minutos cuadrados.

El error absoluto mediano es de **35** minutos.

Lo anterior nos indica que existen outliers que tienen mayor error, influyendo en el error cuadrático.

Modelos Generados



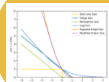
Regresión Lineal

La regresión lineal es una técnica de análisis de datos que predice el valor de datos desconocidos mediante el uso de otro valor de datos relacionado y conocido. Modela matemáticamente la variable desconocida o dependiente y la variable conocida o independiente como una ecuación lineal.



Elastic Net

Es un tipo popular de regresión lineal regularizada que combina dos penalizaciones populares, específicamente las funciones de penalización L1 y L2



SGD Regressor

Modelo lineal ajustado minimizando una pérdida empírica regularizada con SGD. (Stochastic Gradient Descent): el modelo se actualiza a lo largo del camino con un programa de fuerza decreciente (también conocido como tasa de aprendizaje)



Gradient Boosting Regressor

Este estimador construye un modelo aditivo en una forma avanzada por etapas; permite la optimización de funciones de pérdida diferenciables arbitrarias. En cada etapa se ajusta un árbol de regresión sobre el gradiente negativo de la función de pérdida dada.

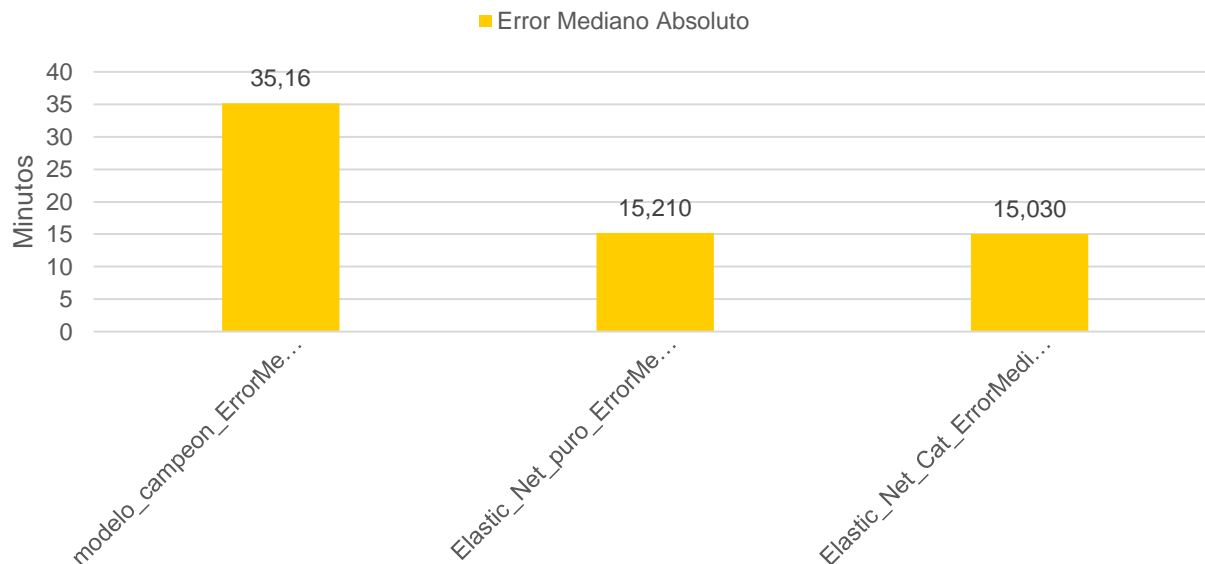


Random Forest Regressor

Un bosque aleatorio es un metaestimador que ajusta una serie de árboles de decisión de clasificación en varias submuestras del conjunto de datos y utiliza el promedio para mejorar la precisión predictiva y controlar el sobreajuste. El tamaño de la submuestra se controla con el parámetro `max_samples` si `bootstrap=True` (predeterminado); de lo contrario, se usa todo el conjunto de datos para construir cada árbol.



Grafica Comparativa de Modelos





Variables que explican el Modelo

ElasticNet

Este modelo de Elasticnet no segmentado, esta en base al total de datos. Dado este modelo las variables que explican el modelos tenemos:
variables temporales de “hora de pedido”, “horario_punta”, “total_product”, “store_branch”, “horario_Valle” y “distancia”

- El error cuadrático medio es 4038.67, mientras que su raíz es 63.55
- El error absoluto mediano es 15.42

ElasticNet Por Categorías < 19 productos

Este modelo de Elasticnet segmentado, tiene las siguientes variables que explican el modelo:
variables temporales de “hora de pedido”, “total_product”, “hora_de_pedido”, “horario_Valle” y “distancia”

- El error cuadrático medio es 1363.61, , mientras que su raíz es 36.93
- El error absoluto mediano es 14.53

ElasticNet Por Categorías > 19 productos

Este modelo de Elasticnet segmentado, tiene las siguientes variables que explican el modelo:
variables temporales de “hora de pedido”, “total_product”, “picking_speed_picker”, y “distancia”

- El error cuadrático medio es 4735.78, mientras que su raíz es 68.82
- El error absoluto mediano es 16.57

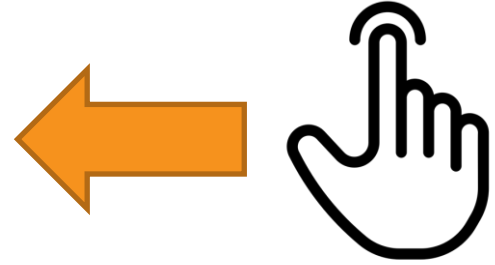


Ejecución Iterativa de Parametros

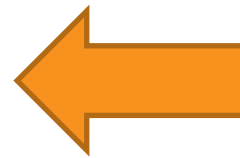



4

Cómo interactuamos









La promesa
Entregas confiables

Entrega

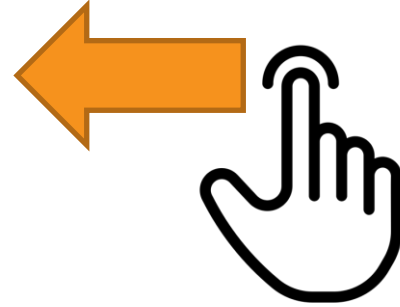
☒

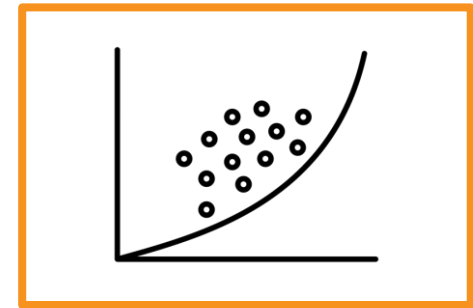
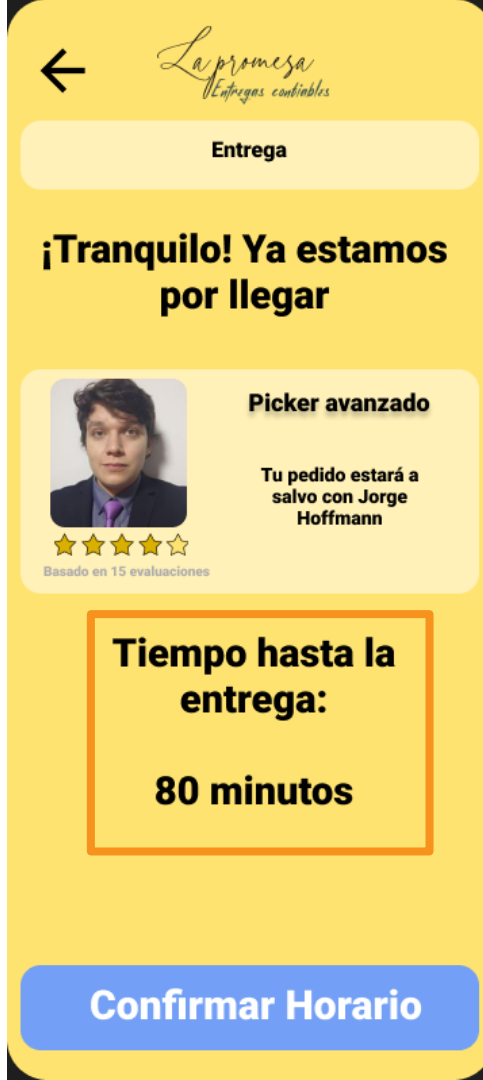
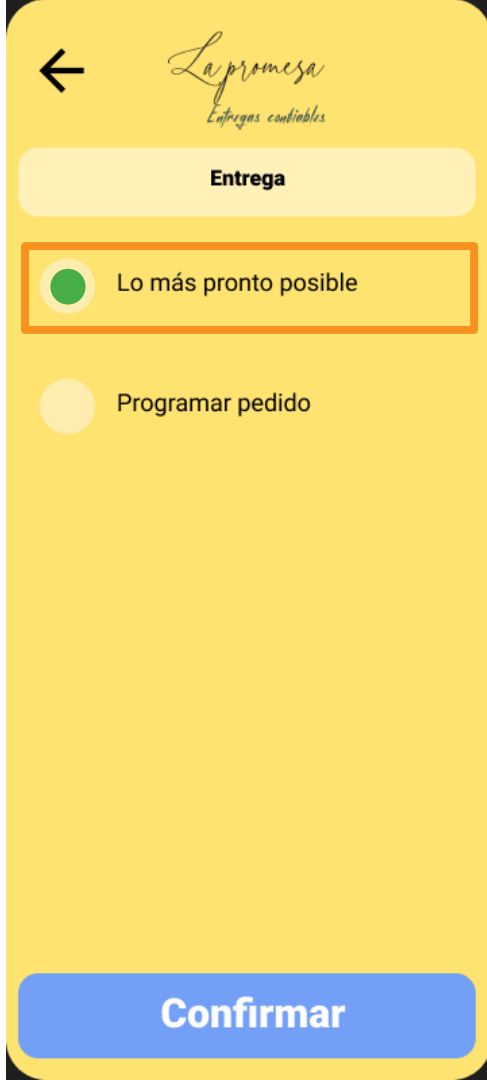
Lo más pronto posible

☐

Programar pedido

Confirmar

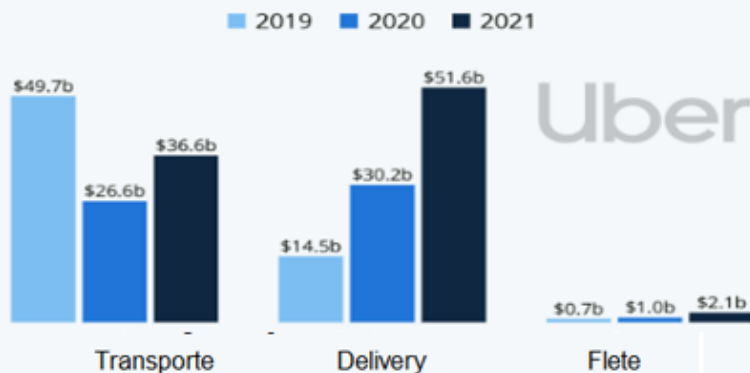




5

Cómo implementar

Uber y la pandemia



Gross bookings are the total dollar value of rides, meal deliveries, and amounts paid by freight shippers without any adjustment for consumer discounts and refunds, driver and restaurant earnings.

Source: Uber

El boom de las entregas continúa

Entrega de paquetes (billones de unidades)

Cambio porcentual (%) respecto a 2020



Parcel volume for B2B, B2C, C2B and consumer consigned shipments with weights up to 31.5kg (70 pounds) across 13 countries.

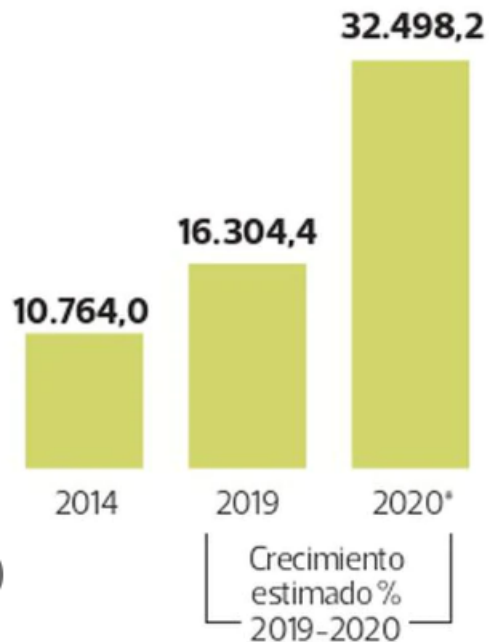
Source: Pitney Bowes

VENTAS POR DELIVERY



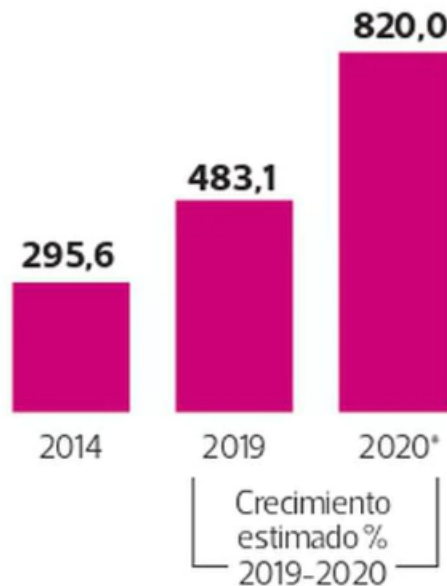
Latinoamérica

Millones US\$, precios corrientes



Chile

Millones US\$, precios corrientes



Fuente: La Tercera, 2021.



Factor crítico y limitante

EL REPARTIDOR



¿Cómo **MOTIVAR**?

¿Por qué los incentivos son importantes?

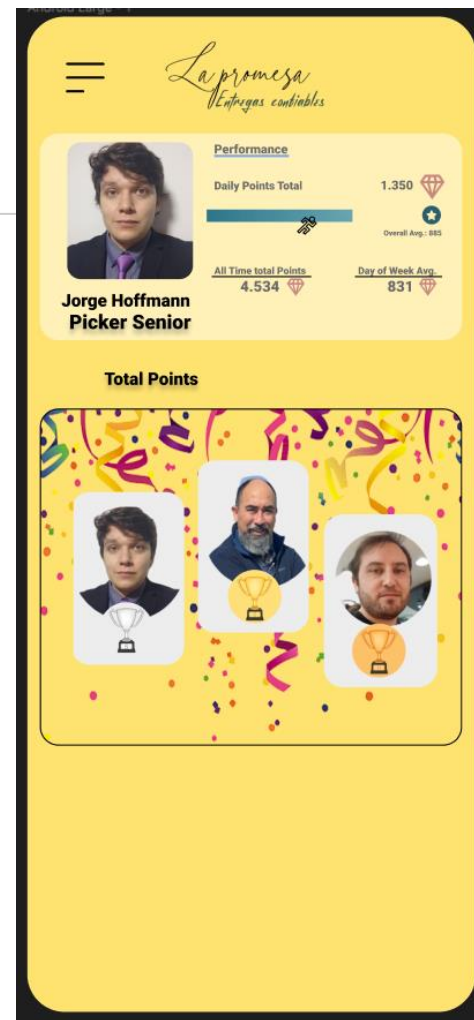
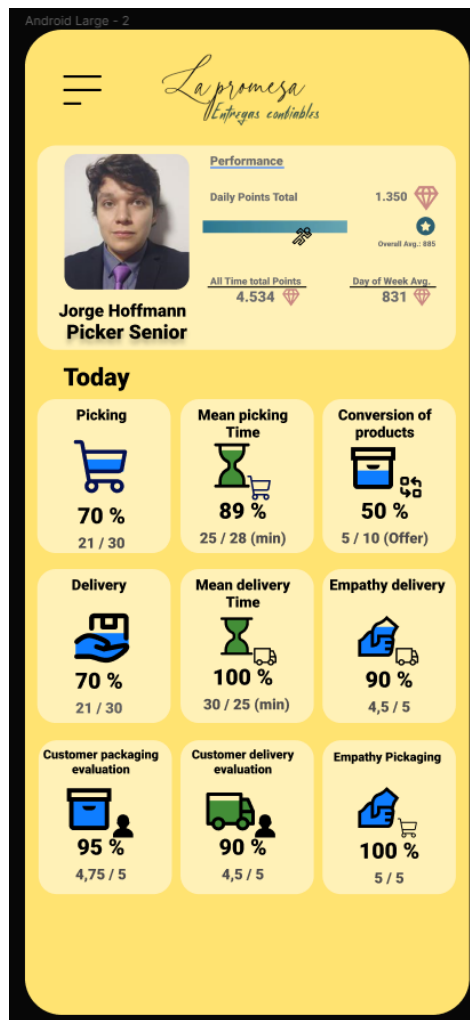
- ⦿ Retención de repartidores
- ⦿ Captación de nuevos repartidores
- ⦿ Incentivos adecuados a cada perfil

Incentivos para repartidores

- Programa de puntos canjeables en variadas categorías como:



- Mayor pago para entregas en horas punta
- Creación de “niveles” o “categorías” de repartidores, que ofrecen flexibilidad horaria para el repartidor
- Alianzas con el picker para disminuir tiempo



6

Conclusiones



Conclusiones



Independiente de los outliers, se sugiere utilizar Elastic Net.



Fue el mejor al ponderar ítems como error, simpleza e interpretabilidad.



Cualquiera sea el camino elegido, debe acompañarse con gestión de personas.