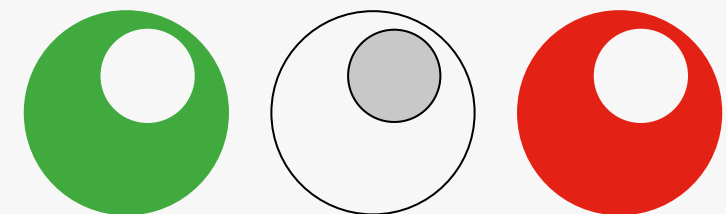


Datanael Cano 

MODELO DE PREDICCIÓN DE VENTA A BORDO

Reto: Viva Aerobus

Damián Jacob Albino Mejía | Julio Enrique Sánchez Guajardo
Victoria Lucero Robles | Francisco Castorena Salazar



OBJETIVO

Crear un **modelo predictivo** (machine learning) para optimizar el **abastecimiento de los productos de venta a bordo** durante el vuelo con el fin de evitar la pérdida de ventas por poco inventario y disminuir las pérdidas de productos en excedente.

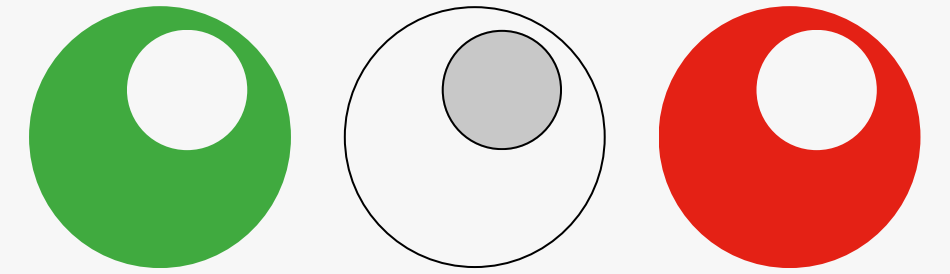
1.1 Modelo predictivo de **pasajeros por vuelo**

1.2 Modelo predictivo de **cantidad de productos por vuelo**

1 - 4 Semanas de pronóstico



BASES DE DATOS



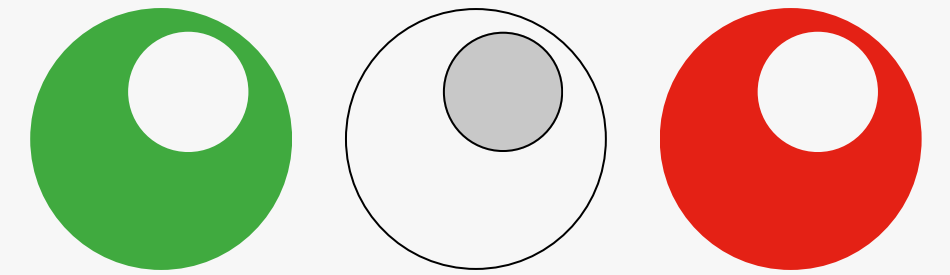
Filghts TEC_Valid.CSV

Flight_ID
Aeronave
DepartureStation
ArrivalStation
Destination_Type
Origin_Type
STD
STA
Capacity
Passengers
Bookings

Sales TEC_Valid.CSV

Flight_ID
ProductType
ProductName
Quantity
TotalSales

PROPUESTA DE VARIABLES



Variables de predicción

Pasajeros

Capacity
Day_of_Week
Hour_of_Day
Month
Destination_Type
DepartureStation
ArrivalStation
Aeronave

Producto

Capacity
Month Day_of_Week
Hour of Day
Duration
Destination_Type
DepartureStation
ArrivalStation
Aeronave
Ocupancia

Nuevos indicadores

Occupancy =
Passengers/Capacity

Duration = STD - STA
Month = [1, ..., 12]
Day_of_Week = [0, ..., 6]
Hour of Day = [1, ..., 24]

METODOLOGÍA

ENTENDIMIENTO
Y LIMPIEZA



PREPARACIÓN
DATA



MODELACIÓN 1



MODELACIÓN 2



VALIDACIÓN



LIMPIEZA:

- VALORES NULL
- SOBREVENTA
- OUTLIERS

PREPARACIÓN DATA:

- NUEVOS INDICADORES
- SEGMENTACIÓN DE LA INFORMACIÓN

MODELOS:

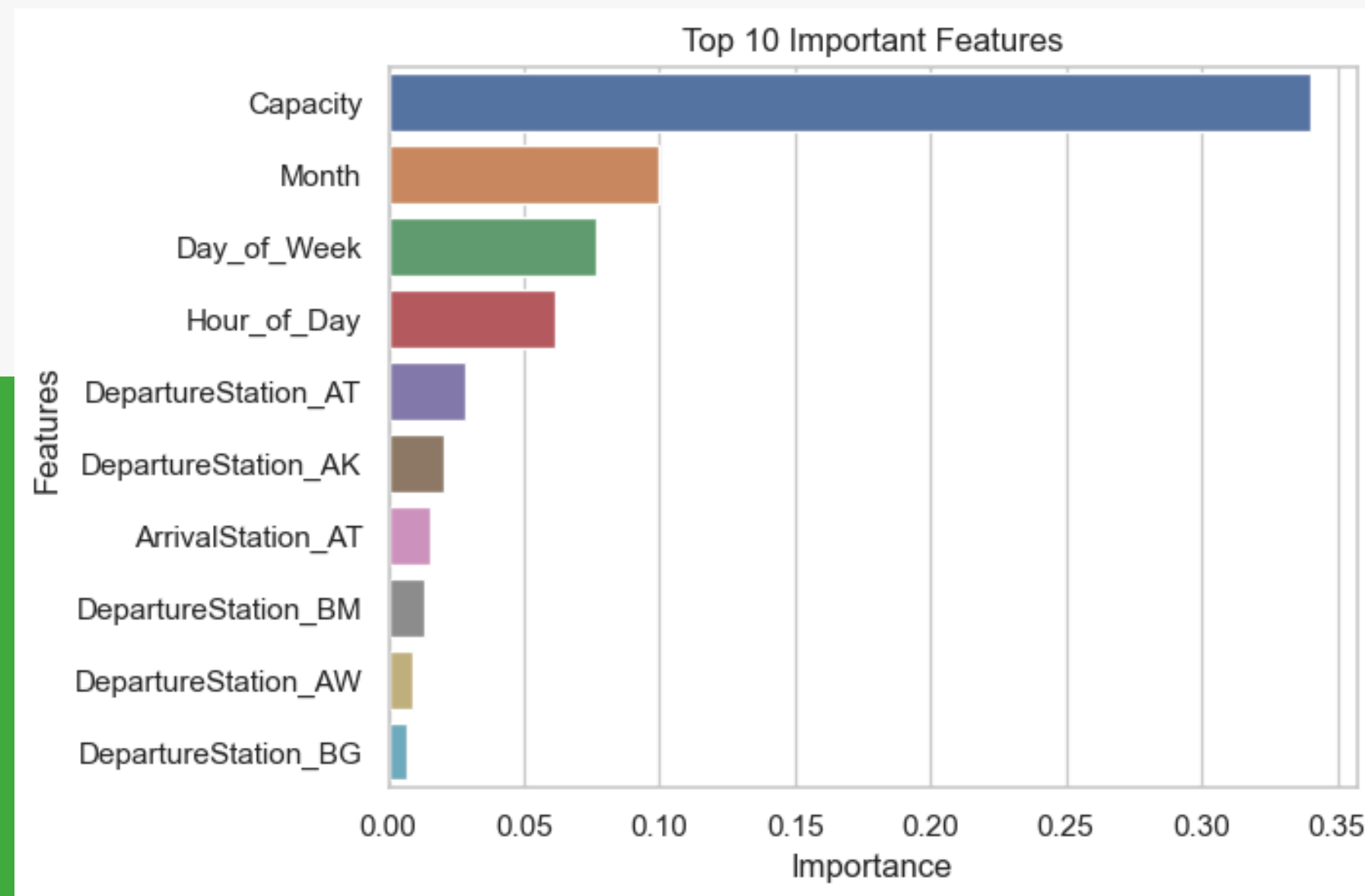
1. PASAJEROS
2. PRODUCTOS

MODELACIÓN DE CLIENTES

RANDOM FOREST

Variables:

```
['Capacity',  
'Day_of_Week',  
'Hour_of_Day',  
'Month',  
'Destination_Type',  
'DepartureStation',  
'ArrivalStation',  
'Aeronave']
```



Parametros:

- `n_estimators=100`

Modelo:

- `RandomForestRegress`

Raíz del error
cuadrático medio
(RMS):

20
pasajeros

Flight_ID	ProductType	ProductName	Predicted Quantity	Actual Quantity
a5a9b15cd551e3f53fdbb8adebc7d13f	Refrescos	Sprite	8	8
eb34a1017b225250a7d58219d1c6dbd1	Botanas	Fritos Limon Y Sal	1	1
89ab5c1469c7656eee308002f0a40f85	Botanas	Ruffles Queso	2	2
e17b4f0cc805a600d730303ef3b86449	Sopas	Nissin Res	7	7
4937b26f5f77bd133939c5d772ce10b7	Botanas	Cheetos	1	1
60c328cc87503662d84d23f25c7fbccc	Licores	Heineken Original	2	2
43c6033bf24dcbba82562e53c28b2d50	Refrescos	Coca Cola Regular	2	2
3be54331ba5975724a05e54d7236978a	Licores	Sol Clamato	1	1
2fe3fdad5f36f1f7c1de5180dedc959f	Galletas	Emperador Chocolate	2	2
b7ffbf8491fe32cb11d895d245e45df4	Refrescos	Agua Natural 600 Ml	8	8
208a5503dffc4e6aa29773de2191bd1	Refrescos	Ciel Mineralizada	1	1

MODELACIÓN CANTIDAD DE PRODUCTOS

RANDOM FOREST

coefficient of
determination
(R²):

0.98

Modelo:

- MultiOutputRegressor
- RandomForestRegressor

Parametros:

- n_estimators=1
- n_jobs=-1

Variables:

['Capacity', 'Month',
'Day_of_Week', 'Hour of
Day', 'Duration',
'Destination_Type',
'DepartureStation',
'ArrivalStation',
'Aeronave', 'Ocupancia']

IMPACTO

- OPTIMIZACIÓN DE LA DEMANDA DE PRODUCTOS A BORDO.
- MEJORA DE LA OCUPACIÓN DE ASIENTOS MEDIANTE PREDICCIONES PRECISAS.
- INCREMENTO DE LA RENTABILIDAD Y LA SATISFACCIÓN DEL CLIENTE A TRAVÉS DE DECISIONES ESTRATÉGICAS BASADAS EN DATOS.