



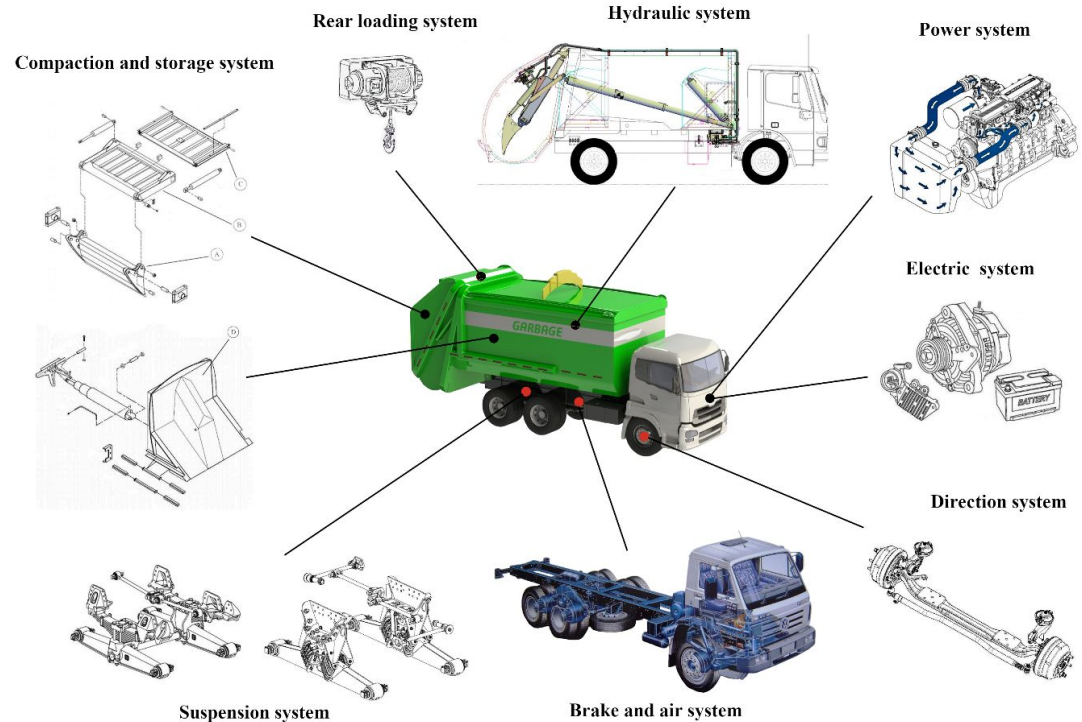
Predicción de sistemas facturados

Sebastián Macías Gallego
Diciembre de 2022
Coding Dojo - Data Science

Datos

Descripción

- **19 categorías para clasificar:**
 - 10 sistemas específicos (Imagen de la derecha más llantas).
 - 2 actividades generales.
 - 1 sistema específico para vehículos de alquiler.
 - 6 sistemas que representan activos de tipo específico.
- Más de 200 activos.
- Dataframe con un total de 17 columnas y 134887 filas.





Problema

Proyectar

Desde el área de mantenimiento es necesario realizar la proyección de costos anual, para ello se requiere conocer la inversión en los diferentes sistemas de los activos.

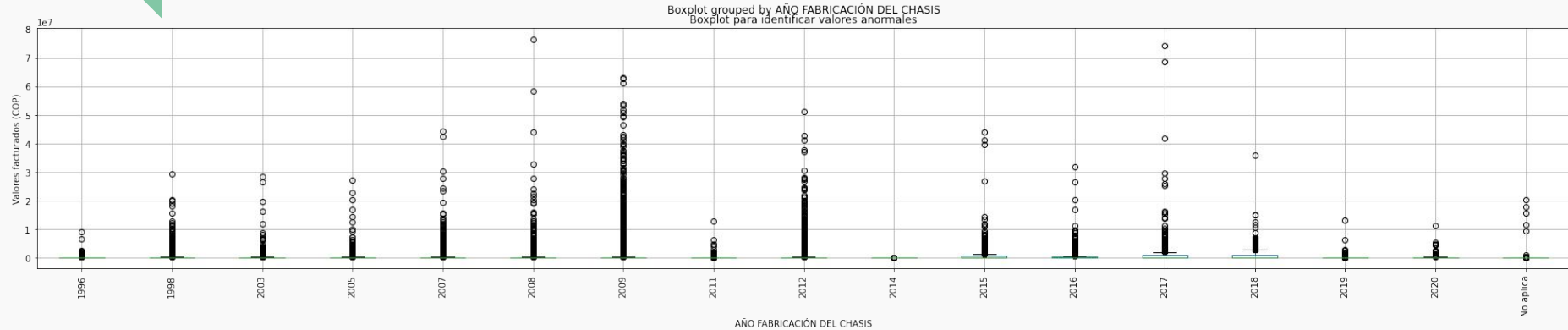
Automatizar

El sistema de facturación actual no permite la clasificación de los sistemas, es una clasificación manual y solo se ha realizado para una parte del historial de facturación.

Clasificar

El reto es asegurar que los sistemas de facturación nueva e histórica sean calificados de manera adecuada.

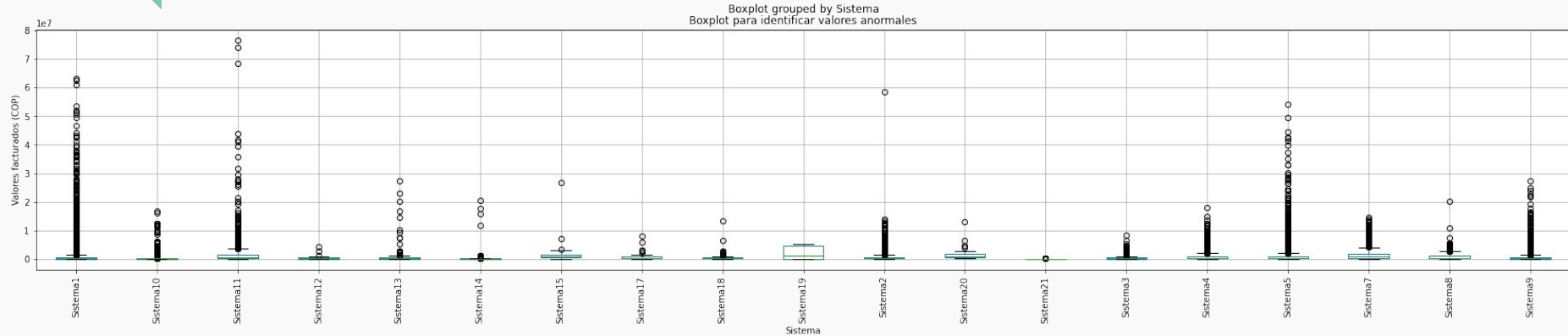
Comportamiento de los datos en relación a activos físicos



Aspectos clave:

1. Los activos físicos se desgastan con el tiempo.
2. Los componentes que fallan deben ser reemplazados, y estos nuevos repuestos pueden ser o no de calidad adecuada.
3. El costo de las reparaciones aumenta con el tiempo para los activos, pues las fallas van siendo cada vez más complejas y requieren tareas de mantenimiento preventivo más complejas hasta procesos de overhaul.

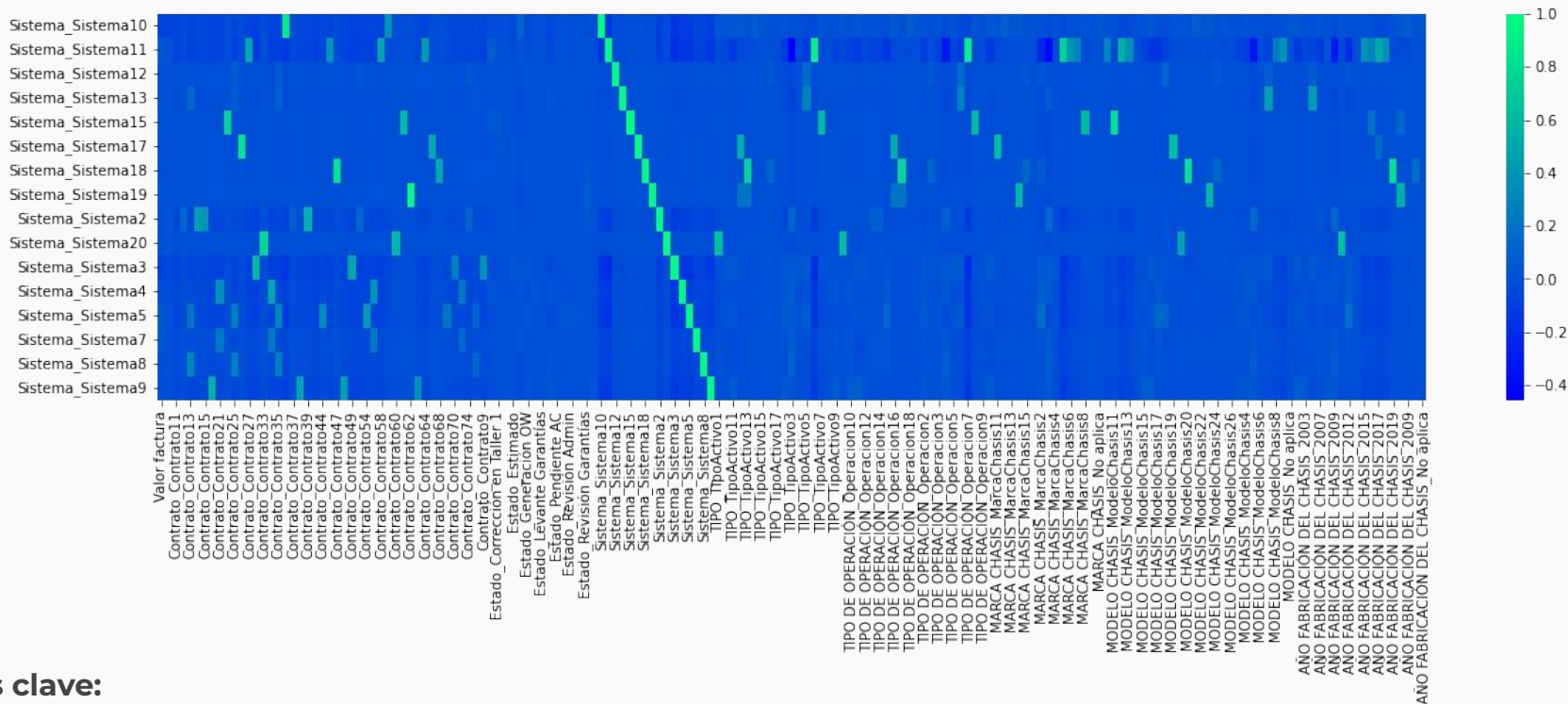
Comportamiento de los datos en relación a activos físicos



Aspectos clave:

1. El cobro de las reparaciones se concentran en ciertos sistemas, los cuales corresponden a componentes principales de los activos, que son el sistema estructural y el motor.
2. La mayor variación en costos está asociada a los sistemas principales, los sistemas asociados a actividades generales permanecen en un rango más acotado, por ejemplo las actividades de lavado y lubricación.

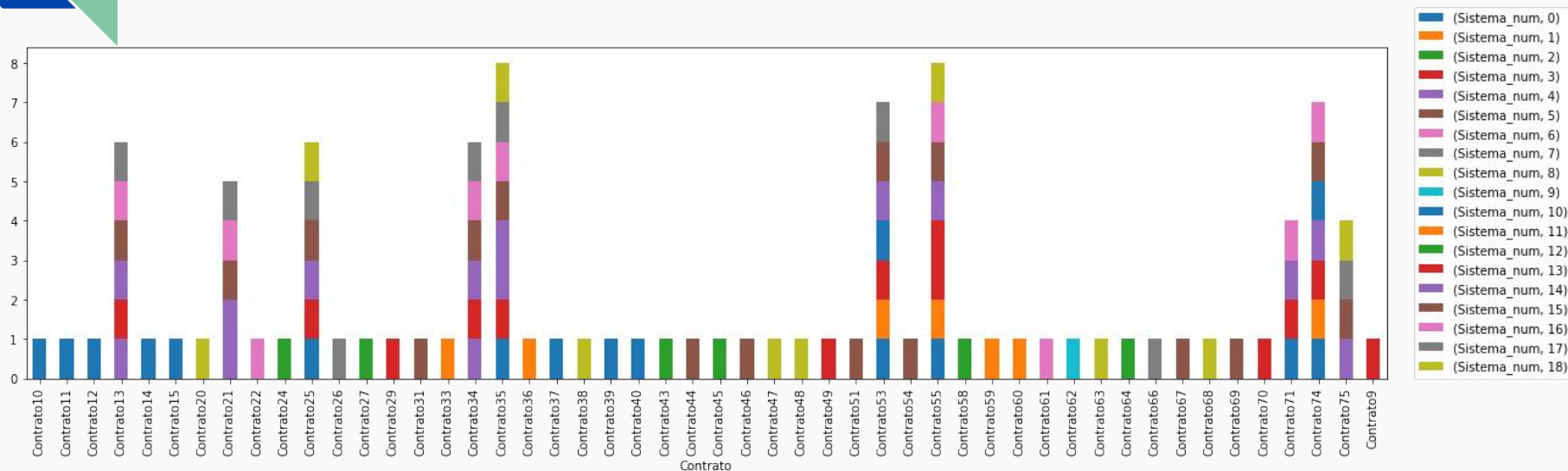
Correlación entre los sistemas y demás variables



Aspectos clave:

Se observa que algunos sistemas tienen una relación importante con el contratista que atiende las novedades y otros sistemas también muestran relación con las características de la operación, el tipo de activo y su año de fabricación.

Relación entre la contratación y los sistemas atendidos



Aspectos clave:

Se observa que un mismo contrato puede atender diferentes tipos de sistemas que normalmente están relacionados, esto es una de las mayores dificultades, poder diferenciar el sistema adecuado cuando es el mismo tipo de activo, la misma operación, el mismo contratista y posiblemente valores facturados similares

Modelos de aprendizaje automático supervisado

Regresión Logística

% de exactitud: 98.5

Tiempo: 266 ms

Sistemas con dificultades en la clasificación (bajo [f1](#) < 0.85):

Sistema 12 - 0.00
Sistema 7 - 0.18
Sistema 8 - 0.77
Sistema 2 - 0.82
Sistema 4 - 0.84

Árbol de decisiones

% de exactitud: 98.3

Tiempo: 46.9 ms

Sistemas con dificultades en la clasificación (bajo [f1](#) < 0.85):

Sistema 12 - 0.00
Sistema 7 - 0.36
Sistema 8 - 0.69
Sistema 2 - 0.83
Sistema 4 - 0.77
*Sistema 13 - 0.76

Bosques aleatorios ***

% de exactitud: 98.1

Tiempo: 766 ms

Sistemas con dificultades en la clasificación (bajo [f1](#) < 0.85):

Sistema 12 - 0.00
Sistema 7 - 0.32
Sistema 8 - 0.72
Sistema 2 - 0.75
Sistema 4 - 0.80

Vecinos más cercanos (KNN)

% de exactitud: 97.6

Tiempo: 3 min 48 s

Sistemas con dificultades en la clasificación (bajo [f1](#) < 0.85):

Sistema 12 - 0.00
Sistema 7 - 0.19
Sistema 8 - 0.70
Sistema 2 - 0.81
Sistema 4 - 0.80



Análisis general de la aplicación de los modelos

Todos los modelos presentan inconvenientes en los mismos sistemas:

- El sistema 12, que corresponde a un sistema de riego de carrotanques, es un sistema exclusivo, pero es atendido por el mismo contratista que atiende los sistemas 5 y 8.
- El sistema 7 y el sistema 4, ambos son sistemas que obran conjunto, uno de ellos representa el sistema de suspensión y el otro el sistema de frenos, ambos atendidos por el mismo contratista y por lo general con reparaciones menores con valores similares y actividades de mantenimiento conjuntas; por tanto, presentan problemas en identificación de ambos sistemas.
- El sistema 2, es un sistema que es atendido a la par con el sistema 1, por el mismo contratista, pero en este caso el valor de las reparaciones del sistema 1 son en muchos casos más elevadas y ocurren con mayor frecuencia, por lo que este sistema si está mucho mejor clasificado.
- El sistema 8, es un sistema que se atiende al tiempo con el sistema 5, uno es dirección y el otro es motor, ambos son cobijados por el mismo contratista, su caso tiene el mismo comportamiento que el sistema 1 y 2.
- El sistema 13, es clasificado de forma adecuada por todos los modelos a excepción del árbol de decisiones.



Aplicación del modelo: **Bosques aleatorios**

El modelo podría ser utilizado con el fin de reducir el tiempo de clasificación de la mayoría de los sistemas, con respecto a las clasificaciones que se realicen de los sistemas 12,7,8, se debe tener un especial cuidado, más aún que el sistema 12 nunca tiene clasificaciones adecuadas; por otro lado, para los sistemas 2 y 4, si bien habría que realizar una revisión cuidadosa, no sería en exhaustiva por las de los sistemas ya mencionados.

Para mejorar la clasificación, sería necesario agregar por ejemplo los ítems asociados a la reparación de esa manera se pueden crear mejores agrupaciones de acuerdo a los elementos que se han cambiado.