

Datos atípicos en las predicciones: una solución al problema

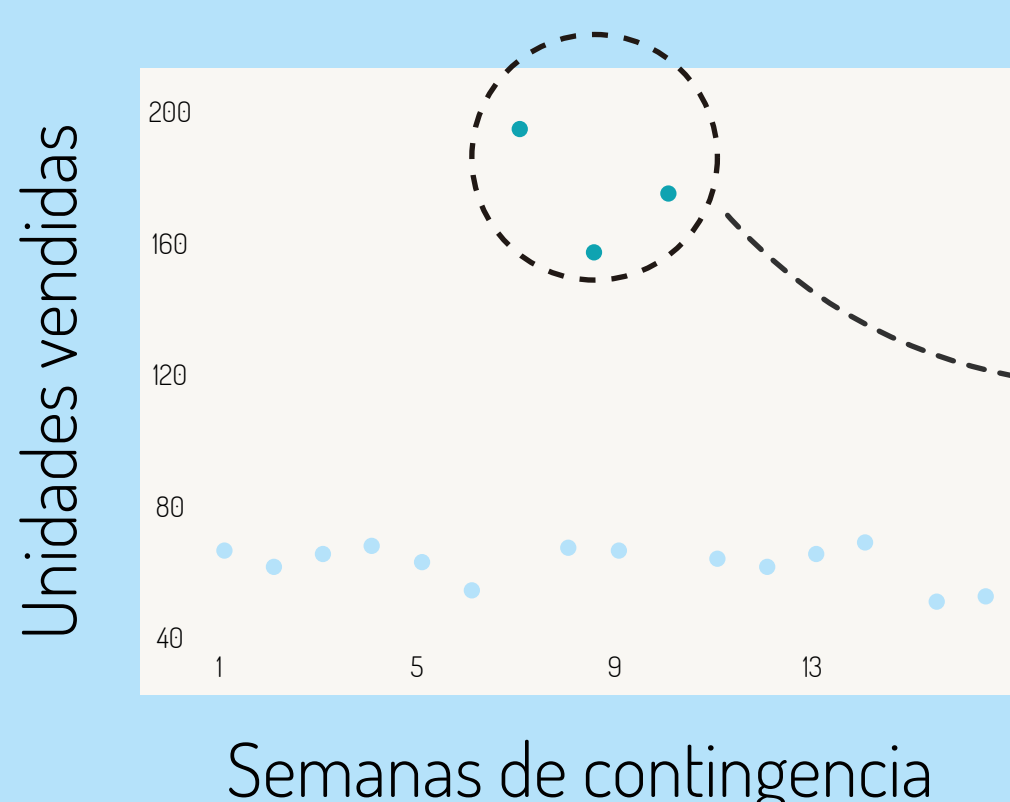
Los estudios estadísticos son importantes para hacer predicciones, pero a veces se encuentran datos que alteran los resultados y hacen que nuestras predicciones no sean correctas. **¿Hay una solución para que estos datos no afecten a los resultados?**

Imaginemos que una tienda quiere predecir el promedio de ventas de un producto desde que empezó la contingencia*

Para eso, utiliza los registros históricos de **cuántas unidades de ese producto se han vendido al día**



*Contingencia producida por el covid-19



La mayoría de datos tiene valores parecidos, y por eso se esparcen por una misma zona

Sin embargo, puede haber algunos valores muy diferentes al resto. **Se conocen como datos atípicos**

Muchas veces, estos surgen por errores (por ejemplo, **cuando sin querer se digita un 0 extra**)

Ante esta problemática, **investigadores de EAFIT mejoraron la segunda etapa de un modelo de aprendizaje estadístico** que ayuda a predecir mejor ante la presencia de datos atípicos

El modelo tiene las siguientes etapas:



Elimina gran parte de los datos atípicos



Disminuye el impacto de datos similares restantes



Establece la relación entre las variables



En esta se reduce el efecto de la información que aún tienen datos atípicos, **y fue optimizada a través de la aplicación de la mediana como una estimación de dispersión muy robusta a datos atípicos**

Este aporte podría representar posibles beneficios para predecir las ventas, demanda de los clientes o portafolios financieros, **sin sezrgarse por datos atípicos**

