

PARTE 1: Industria del Retail

Se debe mencionar en primer lugar que los tipos de impactos son varios, pese a eso por medio de Machine Learning se puede captar: **Patrones de demanda recurrente**, **Decisiones de negocio interna** (Promociones, cambios de precio, localización de los productos), **Factores externos** (Meteorología, arreglos en las calles, tráfico) y **factores desconocidos**, pero que cambian la demanda del cliente. Teniendo eso en mente, se tendría que analizar con que data se dispone. En caso de que no exista, proceder a armar un plan de recolección, siempre y cuando se planifique que datos se buscarán (Posible sesgo).

Por otro lado, cuando hablamos de estimación de demanda con series de tiempo, según la literatura se recomienda hacer uso de aprendizaje automático por medio de Random Forest, Bagging o Boosting en su variación de Adaptive o Gradiente. Estos dos últimos son mecanismo de votación que permite valorizar la fama de un producto o cosa a través de sus variables. Siendo conciso, recomendaría tomar dos caminos sencillos. Con data y sin data.

En caso de no tener data, se recomienda un modelo probabilístico a través de ARMA o ARIMA teniendo las ventas de cada producto versus el tiempo y la cantidad de productos vendido versus el tiempo. Con eso ya se puede obtener un modelo para predecir ventas futuras con estacionalidades.

En el caso de tener data, el espectro se abre. Comenzaría buscando la información del posicionamiento del producto en góndola, ventas en el tiempo, clima asociado a esos días, quiebres de stock, promociones, campañas con descuento, producto retornado, entre otros.

Si nos acercamos a la realidad chilena, en algunos casos los únicos datos que se podrían tener son: Ventas con facturas, ventas con boleta, notas de crédito y notas de débito. Con esto ya se podría preparar una predicción con Stacking, es decir un *Gradient Boosted Tree con suavizamiento exponencial*.

Cabe destacar que se recomendaría seguir el modelo KDD (Knowledge Discovery in Databases), por el cual se debería tener acceso a las bases de datos almacenadas no productivas (cubo, data warehouses u otros).

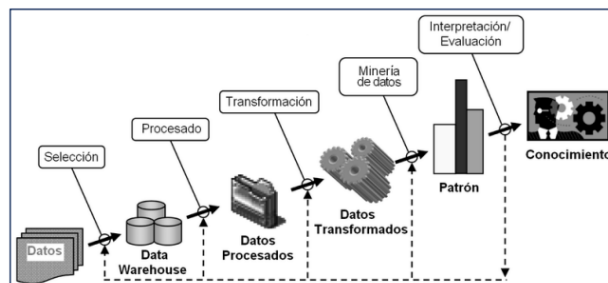


Ilustración 1: Etapas del proceso KDD

Otra mirada menos científica.

- Ver tendencia del hemisferio norte, implicaría conocer las ventas de moda y adelantar la compra el hemisferio sur.
- Se debe considerar también la cantidad de producto que se pueda vender por modalidad de pre-venta (Como el caso actual en la industria automóvil)