

Parte 1

Dada la situación donde un cliente requiere mejorar la toma de decisiones de compras futuras, en primer lugar, acordaría una reunión para conocer al cliente, sus principales actividades, necesidades e inquietudes. En este sentido, podría ocurrir que el cliente crea necesitar un modelo de predicción de demanda cuando en realidad necesita un modelo de gestión de inventario, por lo que necesitamos ratificar el objetivo. Cabe señalar que este objetivo debe especificar también el horizonte del pronóstico y el nivel de agregación.

Una vez presentados los equipos y acordado de manera conjunta que se requiere un modelo de predicción de demanda, asumiendo que éste debe responder a un nivel de agregación semana-tienda-producto con un horizonte de 8 semanas, procedería a realizar un levantamiento de la data del cliente, solicitando las ventas históricas de cada SKU en cada tienda, con un nivel de agregación al menos semanal, con al menos 2 años de historia. Esto va a depender de la frecuencia de venta de los productos y de la cantidad de datos del cliente, por ejemplo, si se venden dos unidades de producto al año, con 2 años es poco probable predecir su venta con precisión, por lo que se asume que los productos tienen una frecuencia de venta relevante y es necesario tratar productos de baja rotación de manera diferente. En este sentido, también solicitaría información de los productos que las empresas de retail suelen manejar, como lo son la jerarquía (categorías/clases/subclases), clasificaciones de rotación, clasificaciones de productos de estación u otra característica que el cliente registre en su maestro de productos. Los datos históricos de promociones y precios de los productos también los solicitaría si es que el cliente los tiene, dado que algunos períodos de ventas históricas podrían estar afectados a aumento de demanda en periodos Cyber, Navidad y Día de la Madre. Adicionalmente, solicitaría la información de stock en el centro de distribución y en tienda, para limpiar periodos en que la venta de un SKU sea cero debido a que el producto estaba con quiebre de stock.

Asumiendo que lo requerido es un pronóstico de demanda semanal con un horizonte de 8 semanas a nivel tienda-SKU, la variable objetivo es continua y se tiene data histórica, por lo que el modelo a elaborar corresponde a uno de tipo supervisado y de regresión. Dependiendo de los datos, yo realizaría distintas iteraciones aumentando el nivel de complejidad del algoritmo de menos a más. Probablemente un modelo de regresión lineal múltiple no sea lo suficientemente robusto y tenga una métrica de error (RMSE o MAE) superior a un modelo Gradient Boosting, SVR o Random Forest Regressor. De este modo, ajustaría distintos tipos de modelos de Machine Learning y, en caso de que ninguno lograra una predicción precisa (bajo R^2 , alto RMSE, alto MAE, alto MAPE), realizaría iteraciones con distintas configuraciones de redes neuronales. Luego de disponer de todos los modelos con sus respectivas métricas, escogería el modelo con mejor ajuste y realizaría un análisis comparativo del resultado del modelo propuesto versus el modelo actual (o estimación del área experta). A través de este análisis, presentando las métricas del modelo propuesto y de la estimación actual, se demuestra ante el cliente en cuánto mejora la precisión, y en cuánto podría aumentar la venta al disminuir el lost sale por quiebre de stock.

Los resultados serían presentados al equipo de trabajo del cliente como un beneficio para su organización y como una herramienta de apoyo para los equipos que se deben encargar actualmente de esa labor, a fin de conseguir adherencia al uso del modelo. Con el visto bueno del cliente, se procede al paso a producción del modelo, más un periodo de seguimiento a convenir.

Parte 2

1.

```
SELECT a.Tienda, COUNT(distinct b.Rut)
FROM POS_TRX a LEFT JOIN POS_TRX_CLIENTE b on a.ID_TRX = b.ID_TRX
WHERE a.Tipo_transaccion = 15
GROUP BY a.Tienda;
```

2.

```
SELECT YEAR(b.Fecha_transaccion) as anio, a.Tienda, SUM(b.Monto_pago) as venta_anual
FROM POS_TRX a LEFT JOIN POS_TRX_PAGO b on a.ID_TRX = b.ID_TRX
GROUP BY YEAR(b.Fecha_transaccion), a.Tienda
ORDER BY anio desc, venta_anual desc;
```

3.

```
SELECT
    YEAR(b.Fecha_transaccion) as anio,
    a.Tienda,
    a.Tipo_transaccion,
    AVG(b.Monto_pago) as monto_promedio_transaccion
FROM POS_TRX a LEFT JOIN POS_TRX_PAGO b on a.ID_TRX = b.ID_TRX
GROUP BY YEAR(b.Fecha_transaccion), a.Tienda, a.Tipo_transaccion
ORDER BY anio desc, AVG(b.Monto_pago) desc;
```

4.

```
SELECT Medio_de_pago from (
    SELECT
        Medio_de_pago,
        SUM(Monto_pago) as monto_total_medio_pago
    FROM POS_TRX_PAGO
    GROUP BY Medio_de_pago
    ORDER BY SUM(Monto_pago) desc)
LIMIT 1;
```

Se asume como principal medio de pago el medio por el cual se acumula mayor monto pagado