










<h2>NUEVOS DATOS Y REENTRENAMIENTO</h2>  <p>El modelo podrá ser re-entrenado al final de cada mes. Los clientes pueden cambiar de cluster a lo largo del tiempo.</p>	<h2>PREDICCIÓN ON/OFF</h2>  <p>Modelo en batch. Reentrenado de manera mensual. Se entregará una vez al mes. No entregamos una predicción si no una fotografía de la segmentación a fecha de la entrega. No es estática. De manera anual, según la evolución se podrá entregar una nueva propuesta en función de si aparecieran nuevos grupos o fuese necesario un nivel de detalle más alto.</p>	<h2>PROPUESTA DE VALOR</h2>  <p>Mejora de la segmentación actual del grupo para la orientación de la actividad comercial.</p> <p>Actualmente cuenta con 3 segmentos sin seguimiento, estáticos.</p> <p>Petición directora comercial, 7-8 grupos.</p>	<h2>ORÍGENES DE DATOS</h2>  <p>Partimos de 3 datasets:</p> <p>Productos, Sociodemográfico, Commercial Activity</p> <p>Datos facilitados por el departamento de IT. Total 17 particiones. Inicio: 28.01.2018 hasta el 28.05.2019</p> <p>Formato CSV.</p>	<h2>TAREA DE ML</h2>  <p>Aprendizaje no Supervisado, pero guiado por el data scientist.</p> <p>Tras el análisis de negocio.</p> <p>Tarea: Clustering (agrupamiento)</p> <p>Se utiliza el algoritmo KMeans.</p>
<h2>EVALUACIÓN EN SERVICIO Y ALM</h2>  <p>Se tendrá que realizar una medición del número de clusters óptimos a través del diagrama del codo para validar posibles cambios en el comportamiento de nuestros clientes causado por la incorporación de nuevos productos y servicios o debido a causas ajenas a la empresa.</p>	<h2>MÉTRICA DE EVALUACIÓN</h2>  <p>Ajustamos el número de Clusters óptimos mediante la inercia/dispersión de los distintos clusters.</p> <p>Se utiliza el diagrama del codo.</p>		<h2>ATRIBUTOS</h2>  <p>Los atributos están previamente definidos en el txt entregado por IT. Se generan los siguientes atributos adicionales:</p> <ul style="list-style-type: none"> • Número de altas • Número de bajas • Número de cobros • Permanencia por tipo de producto. (familias de producto: financiación, cuentas, inversión...) <p>Se exploran otros descartados finalmente que no entrarán en el modelo (es español, region code).</p>	<h2>DEFINICIÓN DEL UNIVERSO Y OBJETIVO</h2>  <p>El objetivo es asignar un cluster a cada cliente. El identificador de cliente se repite a lo largo de las particiones según antigüedad. Necesitamos agrupar por pk_cid para obtener los datos de manera que cada fila corresponde a un cliente.</p> <p>Población: En todo el histórico tenemos un total de 105.989 clientes únicos que no han contratado nunca ningún producto. Contamos con un total de 350.384 clientes únicos que han contratado algún producto</p> <p>Los clientes que nunca han tenido un producto activo, son clientes en Base de datos que hemos agrupado previamente como clientes en hibernación y no forman parte del grupo de clientes sobre el que hacemos el clustering.</p>
	<h2>USO DEL MODELO, TOMA DE DECISIONES I EXPLICABILIDAD</h2>	<p>La entrega de los resultados se hará mediante una ficha de cliente, explicará la relación cliente-producto. Se entregan visualizaciones gráficas indicando las características de consumo y demográficas de los grupos de clientes.</p>	<p>Los grupos de clientes no serán estáticos, los grupos de clientes pueden evolucionar según el comportamiento en el tiempo del cliente-producto. El modelo asistirá en las decisiones de la actividad comercial.</p>	