#### **1. Introducción**

**Contexto del Proyecto:** La empresa GHA está interesada en estimar el valor de compra semanal de mercadería para mejorar la planificación y optimización de sus inventarios y compras. Para ello, se ha desarrollado un modelo predictivo utilizando técnicas de Machine Learning y se ha creado un dashboard que proyecta las compras para el año 2024 basado en los datos de 2023.

**Objetivos del Proyecto:**

* Desarrollar un modelo predictivo que estime el valor de compra semanal de mercadería regulares.
* Crear un dashboard que proyecte el valor de compra para el año 2024, considerando un aumento del 25% en variables clave durante semanas específicas.

**KPIs Definidos:**

* Valor de Compra semanal
* Cantidad de Sacos semanal
* Cantidad en Peso semanal

**Métricas de Negocio:**

* MAE, RMSE, y R² del modelo predictivo
* Incremento porcentual en las proyecciones para 2024

**2. Análisis de Machine Learning**

**Descripción de los Datos:** La base de datos utilizada contiene información de las compras realizadas por la empresa, incluyendo variables como 'Nombre del Agricultor', 'Fecha de Liquidación', 'Producto', 'Variedad', 'Calidad', 'Cantidad de Sacos', 'Cantidad en Peso', 'Valor de Compra', 'Año', 'Lugar', y 'Semana'.

**Preprocesamiento de Datos:**

* **Limpieza de Datos:** Se eliminaron registros incompletos y outliers para mejorar la precisión del modelo.
* **Transformaciones y Normalización:** Se realizaron transformaciones de datos para normalizar las variables y facilitar el entrenamiento del modelo.

**Selección de Características:** Se seleccionaron variables clave como 'Cantidad de Sacos', 'Cantidad en Peso', 'Variedad', 'Calidad', 'Semana', y 'Lugar' para construir el modelo predictivo.

**Desarrollo del Modelo Predictivo:**

* **Algoritmo Utilizado:** Se utilizó un algoritmo de regresión lineal para estimar el valor de compra semanal.
* **Parámetros y Configuración:** Se ajustaron los parámetros del modelo para optimizar su rendimiento.
* **Métricas de Evaluación:** Se usaron métricas como el error cuadrático medio (MSE), error absoluto medio (MAE) y el coeficiente de determinación (R²) para evaluar y validar el modelo.

**Resultados del Modelo:** Los resultados mostraron un MAE inferior a 1000 y un RMSE inferior a 1500, métricas que la empresa considera aceptables. Además, el modelo obtuvo un coeficiente de determinación (R²) del 85%, lo que indica un buen ajuste y confiabilidad del modelo.

**3. Desarrollo del Dashboard**

**Descripción de los Datos:** Se utilizó la misma base de datos para el desarrollo del dashboard, enfocándose en las proyecciones para el año 2024.

**Proyección de Datos:**

* **Metodología:** Se aplicó un incremento del 25% en las variables 'Cantidad de Sacos', 'Cantidad en Peso' y 'Valor de Compra' para las semanas 19, 20, 21 y 22 del año 2024.
* **Justificación:** Estas semanas fueron seleccionadas debido a su alto rendimiento histórico en ventas.

**Implementación del Dashboard:**

* **Herramientas Utilizadas:** Se utilizó Power BI para crear el dashboard interactivo.
* **Descripción de las Visualizaciones:** El dashboard incluye visualizaciones clave como la suma del valor de compra por año, cantidad de sacos por variedad, calidad y producto, y un mapa que muestra las compras por lugar.

**Resultados y Análisis:** El dashboard proyecta un aumento significativo en el valor de compra para las semanas seleccionadas en 2024, basándose en un incremento del 25% en las variables clave. Esto ayudará a la empresa a tomar decisiones informadas sobre sus compras futuras, optimizando el proceso de adquisición y mejorando la gestión del inventario. La visualización interactiva facilita la identificación de tendencias y patrones en los datos, permitiendo ajustes rápidos y eficientes en las estrategias de compra.

**4. Reflexión sobre los Aprendizajes:**

* El análisis y el proceso de machine learning permitieron obtener una comprensión profunda de los datos de compra de la empresa.
* El uso de Ridge Regression resultó efectivo para manejar la multicolinealidad y proporcionar predicciones precisas.
* Los dashboards fueron efectivos para comunicar los resultados del análisis de manera clara y concisa.
* La interactividad y las visualizaciones facilitaron la comprensión de las proyecciones y las tendencias.