



**TRABAJO DE GRADO**  
**Opción Seminario-Diplomado.**

**ESTRATEGIA COMPUTACIONAL PARA PREDECIR ESTIMACION DE DEMANDA**  
**EN ABASTECIMIENTO AUTOMATICO, UTILIZANDO ALGORITMOS DE MACHINE**  
**LEARNING**

Corporación Universitaria Remington.

Ingenierías

Especialización en dirección de operaciones en mejoramiento  
continuo.

Estudiante:

Deibys Rafal Bautista Suarez

Tutor: Juan Carlos Briñez de León

Opción de Trabajo de grado Seminario-Diplomado.

2025.

## **Dedicatoria**

A mis hijas, son mi motor para cada reto, para afrontar cada oportunidad, aprender y mejorar como padre y persona. A mis padres, especialmente a mi madre, por cada consejo y enseñanza dada en cada paso de mi vida, es mi ángel.

## **Agradecimientos**

A DIOS, A LA UNIVERSIDAD. AL PROFE DEL SEMINARIO, POR SU ESPERA; y a todas aquellas personas que suman en mi proceso de formación.

**Contenido**

Resumen.....	5
Palabras clave.....	6
Marco conceptual y contextual .....	7
Contexto .....	7
Objetivos .....	12
Objetivo general.....	12
Objetivos específicos.....	12
Desarrollo e implementación del aprendizaje .....	13
Preparación y análisis de los datos.....	14
Modelo de toma de decisiones .....	15
Análisis de desempeño .....	16
Validación del modelo .....	17
Conclusiones y trabajos futuros .....	18
Referencias bibliográficas .....	19

## Resumen

Este trabajo de grado tiene como objetivo aplicar técnicas de análisis de compras de fábricas con la inminente necesidad de solucionar los recurrentes problemas de abastecimiento que afectan la producción con el objetivo de predecir la cantidad y los tiempos de abastecimiento la materia prima necesarios para un año de producción. Bajo esto, se utilizó un conjunto de datos con información de año de compra de artículos, incluyendo variables como fecha de compra, proveedor descripción artículo, número de orden, entre otros.

A lo largo del proyecto, se realizó un análisis completo que incluyó la limpieza de los datos, transformación de variables y selección de características relevantes. Posteriormente, se entrenaron distintos algoritmos de clasificación supervisada como regresión logística, árboles de decisión, Random Forest, SVM, entre otros, con el fin de predecir si un préstamo debía ser aprobado o rechazado.

Este conjunto de datos es una instantánea de un panorama minorista ficticio, que captura los atributos esenciales que impulsan las operaciones minoristas y las interacciones con los clientes. Incluye detalles clave como el ID de la transacción, la fecha, el ID del cliente, el género, la edad, la categoría del producto, la cantidad, el precio unitario y el importe total. Estos atributos permiten una exploración multifacética de las tendencias de ventas, las influencias demográficas y los comportamientos de compra.

**Palabras clave**

Machine Learning; Abastecimiento; Artículo; Orden de Compra; Algoritmos de predicción; Análisis de datos; Modelos de decisión; Regresión logística.

## Marco conceptual y contextual

### Contexto:

#### Modelos de Abastecimiento

En la actualidad los modelos de abastecimiento automáticos guardan una relación mutua con machine learning, ya que, a través de estos, el sistema logístico de inventarios, predicción y abastecimiento dentro de la cadena de suministro está cambiando la manera en la que distintas compañías y fábricas usan la información para la precisión en la toma de decisiones. Así mismo, las técnicas de aplicación para dar solución a los problemas complejos son a través del proceso analítico jerárquico o herramienta AHP, por lo cual las características de las matrices de criterios y datos darán resultados con la mejor decisión para las organizaciones. De acuerdo con Girlich y Chikán [6], el desarrollo conjunto de las teorías de inventarios y la aplicación de las matemáticas y la estadística se inició desde los años 50 cuando la Oficina de Investigación Naval de California destinaron recursos para la investigación en el área.

Más recientemente, autores como Snyder et al. [14] han utilizado la simulación como técnica de mejoramiento de las metodologías de pronósticos mediante suavización exponencial. Los autores hacen una diferencia entre los ítems que tienen una alta rotación y los de baja rotación (caso que se presenta comúnmente en la mayoría de las industrias) y sistematizan el análisis de series de tiempo, a través del uso de herramientas de simulación como la Montecarlo para predecir los niveles de demanda y su relación con los Lead Times.

Este trabajo de grado propone aplicar la evidencia de una data una fábrica de producción del año 2020 al 2023, con el fin de predecir la cantidad y los tiempos de abastecimiento de los accesorios que serán necesarios para un año de producción. Esto facilitará la toma de decisiones en el área de compras de la empresa, ya que, se conocerá con anticipación lo que será utilizado, sin generar compras excesivas que generen stocks inmovilizados para determinados artículos.

Además, permitirá una mejor negociación con su portafolio de proveedores, porque se sabrá con exactitud las cantidades a solicitar, costo unitario y el tiempo de despacho de acuerdo a flujo logístico.

### **Algoritmos de Machine learning en Modelos de Abastecimiento**

Distintos analistas y expertos en logística han estudiado la criticidad y beneficios que ofrecen los algoritmos de machine. El machine learning o aprendizaje automático es un nuevo tipo de computación. A diferencia de la tradicional, donde se programan algoritmos estáticos, con machine learning se puede procesar un enorme volumen de información e identificar patrones. A partir de la repetición continua de este análisis, el propio algoritmo perfecciona su funcionamiento continuamente hasta obtener resultados cada vez más precisos. Aquí se emplean los siguientes tipos de algoritmos de machine:

- **Aprendizaje supervisado:** es el tipo de machine learning más extendido. Estos algoritmos funcionan con datos de entrada y de salida ya clasificados

de antemano. Por ejemplo, si se está buscando que pueda distinguir entre imágenes de peras y manzanas, el desarrollador señalará en qué fotos aparece una pera y en cuál una manzana. Tras analizar miles de fotos, el algoritmo aprenderá a diferenciarlas.

- **Aprendizaje no supervisado:** este tipo de algoritmo de machine learning no cuenta con el etiquetado previo de los datos de entrada ni de salida. Es el propio sistema el que tiene que analizar todo el conjunto de datos y tratar de detectar patrones basándose en similitudes para etiquetarlos sin la ayuda de una persona.
- **Aprendizaje por refuerzo:** se sitúa en un terreno intermedio entre los anteriores. En este caso, el algoritmo aprende a base de ensayo-error y ajusta sus acciones de forma dinámica en función del feedback que recibe del mundo exterior.

En este trabajo se pretende aplicar y comparar varios de estos algoritmos para resolver un problema real: estimar la probabilidad de que un préstamo sea aprobado, a partir de los datos de solicitantes registrados en una entidad financiera. El conjunto de datos contiene variables como género, casado, dependientes, ingresos, entre otros.

## **1.2 Descripción de caso de estudio.**

El caso se enfoca en el análisis y desarrollo de un modelo de abastecimiento de demanda basada en algoritmos de machine learning, con el objetivo de predecirla exactitud de demanda de un producto de acuerdo al stock de

inventario sin que haya quiebre en los tiempos de espera. Para ello, se empleará un conjunto de datos de demanda de producto en un periodo de 15 meses, desde el año 2024 al año 2025.

Este tipo de demanda a predicción es utilizado en los modelos de simulación de producción, en los modelos de reabastecimiento automático y es fundamental para mantener Kardex de inventarios oxigenados, detallando atributos a los insumos y productos, como también la antigüedad de rotación de los mismos. Bajos este modelo se demostrará con datos y hechos la calidad panel de los inventarios, antigüedad de productos, demanda predictiva exacta y las estimaciones del método de paretto.

**1.3 Pregunta problema:**

¿Cómo desarrollar una estrategia computacional para estimar la demanda de productos a partir de los datos de los atributos de este, utilizando algoritmos de aprendizaje supervisado de machine learning?

**1.4 Hipótesis:**

El análisis computacional de los datos de datos de demanda por modelos de simulación, utilizando algoritmos de aprendizaje supervisado, permitirá predecir con precisión de pronóstico en cantidades y tiempo, manteniendo así eficiencia en los resultados esperados con colchones de seguridad máximo y mínimo.

## Objetivos

### Objetivo general.

Implementar una estrategia computacional para predecir la exactitud de la demanda bajo, utilizando algoritmos de aprendizaje supervisado de machine learning.

### Objetivos específicos.

- ✓ Caracterizar y procesar los datos de interés, con miras a la toma de decisiones informadas.
- ✓ Implementar un algoritmo de Machine Learning para la predicción de demanda de productos en fábrica.
- ✓ Evaluar y analizar el modelo de reabastecimiento con los algoritmos implementados en la demanda estimada.
- ✓ Validar el correcto funcionamiento del modelo de reabastecimiento automático.

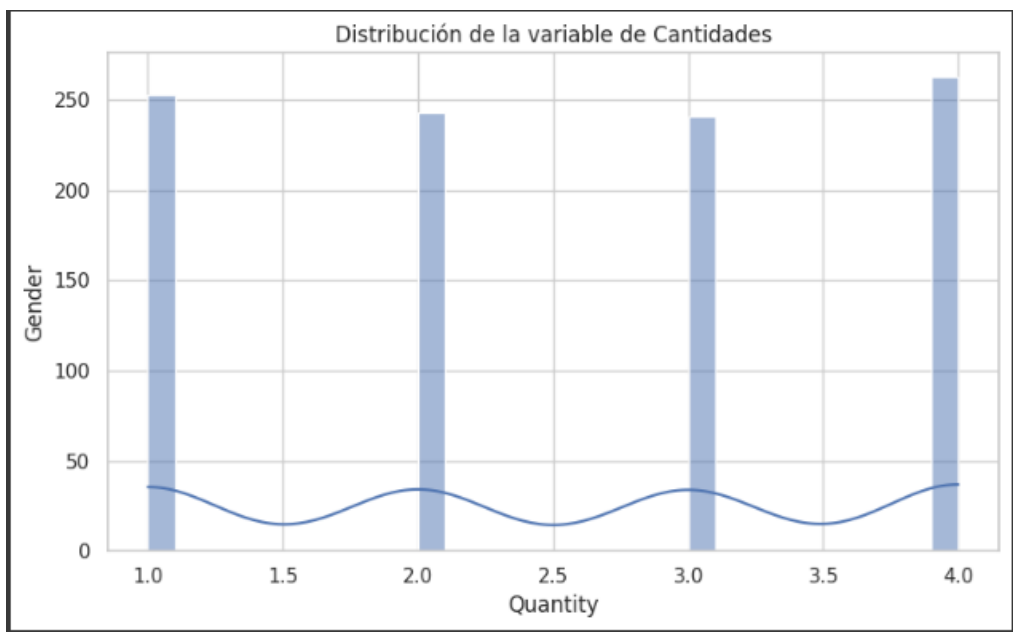
## Desarrollo e implementación del aprendizaje

se diseñó una estrategia computacional basada en algoritmos de machine learning, específicamente del tipo aprendizaje supervisado orientado a clasificación binaria. Esta metodología fue aplicada sobre un conjunto de datos de demanda de productos que simula los atributos de seguridad máximo y mínimo, como también la data de ventas según el ID del producto.

1. **ID de transacción:** un identificador único para cada transacción, que permite el seguimiento y la referencia.
2. **Fecha:** la fecha en que se realizó la transacción, que proporciona información sobre las tendencias de ventas a lo largo del tiempo.
3. **ID de cliente:** un identificador único para cada cliente, que permite un análisis centrado en el cliente.
4. **Género:** el género del cliente (hombre/mujer), que ofrece información sobre patrones de compra basados en el género.
5. **Edad:** La edad del cliente, lo que facilita la segmentación y la exploración de las influencias relacionadas con la edad.
6. **Categoría de producto:** la categoría del producto adquirido (por ejemplo, Electrónica, Ropa, Belleza), que ayuda a comprender las preferencias de productos.
7. **Cantidad:** El número de unidades del producto comprado, lo que contribuye a obtener información sobre los volúmenes de compra.
8. **Precio por unidad:** El precio de una unidad del producto, que ayuda en los cálculos relacionados con el gasto total.
9. **Importe total:** el valor monetario total de la transacción, que muestra el impacto financiero de cada compra.

## Preparación y análisis de los datos

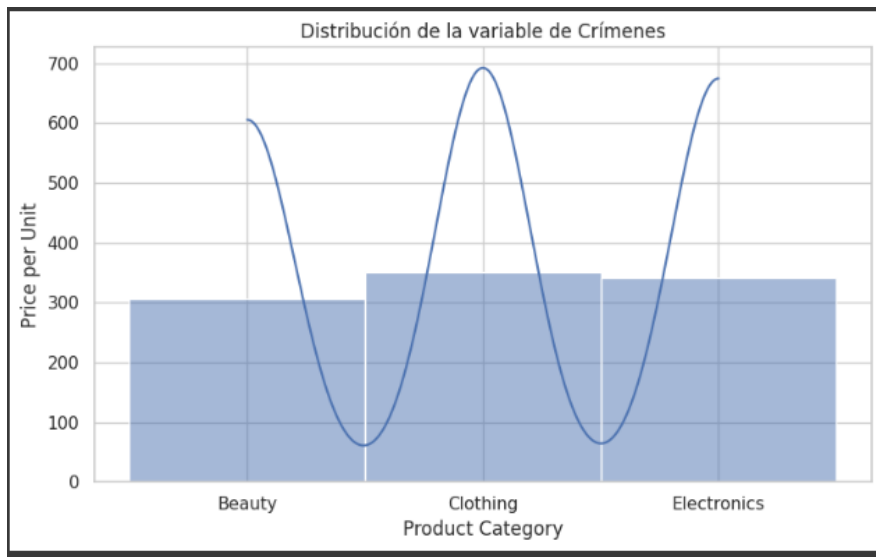
Para iniciar con un análisis preliminar de los datos, se han generado dos gráficos básicos: la distribución de la variable de salida y en procesamiento de datos ETL de las categorías. A continuación, se presentan los gráficos obtenidos en Google Colab. EL gráfico indica que las cantidades solicitadas de ciertas referencias según Su Género para hacer demanda de productos, presentan una distribución importante en la rotación periódica de los productos, ya que esto ayuda a tener una población objetivo y focalizada en las necesidades



**Grafico 1**

### **Análisis variables demanda**

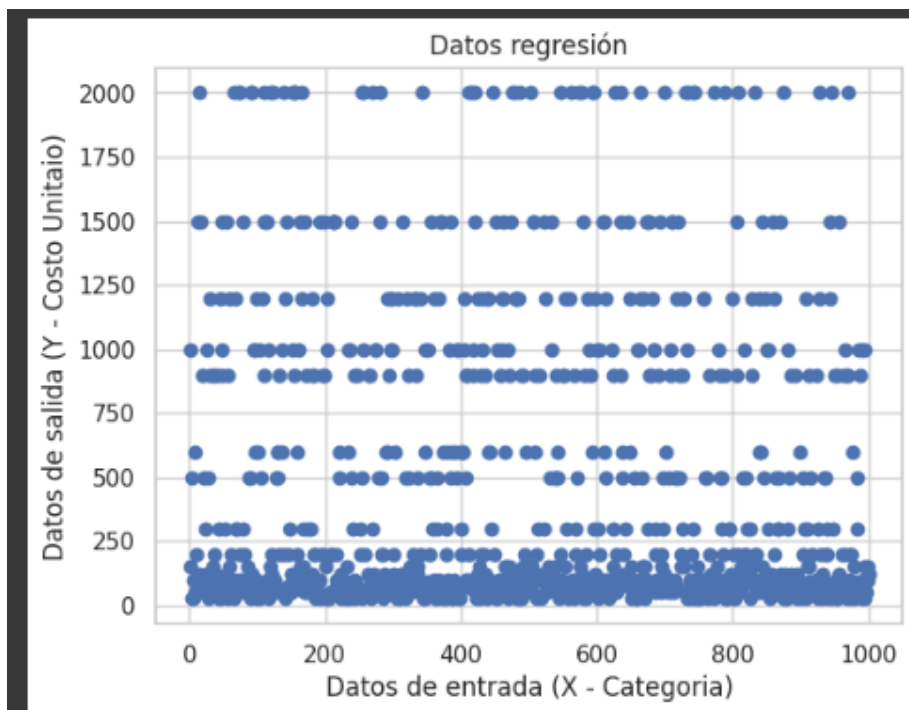
Aplicando distribución de datos de salida d precio unitario vs categoría, existe una relación fuerte en los precios menos a 100 pesos, ya que hay una concentración alta de categorías de productos Consumidos por el cliente que preponderan en la demanda y reabastecimiento de la cadena de suministro



**Gráfico 2**

**Datos de regresión:**

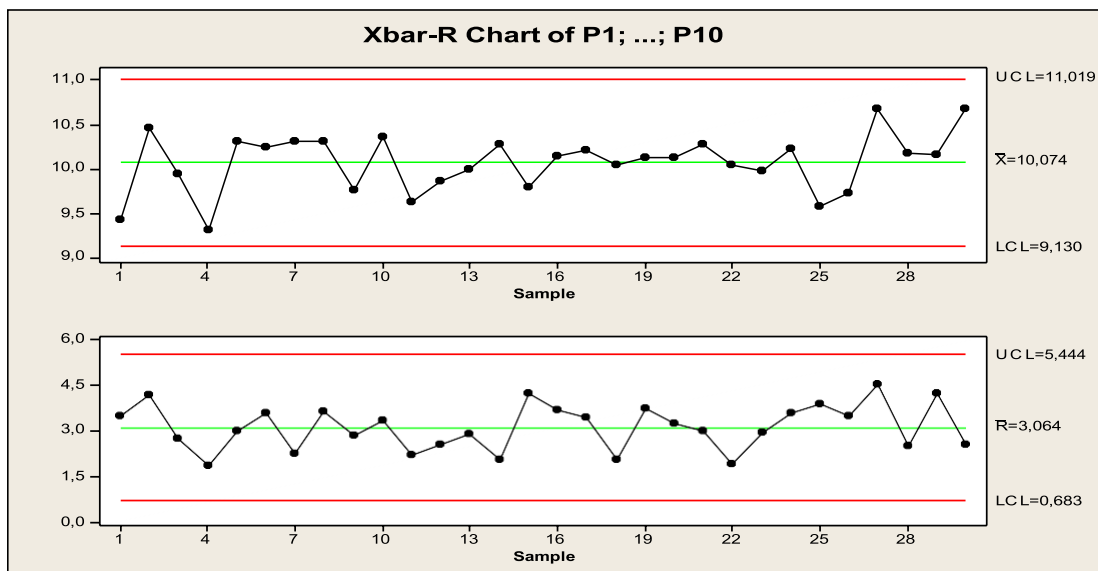
Para los atributos a de categoría de productos, los dispositivos electrónicos con costo inferior a 250 peso tienen una alta concentración de demanda vs aquellos que en el paretto tiene un costo superior a 1000 pesos. Bajo estos datos de región la demanda de categorías de productos resulta ser un poco tediosa de analizar, ya que hay datos dispersos con una sola orientación.



**Gráfico 3**

### ***Control Estadístico***

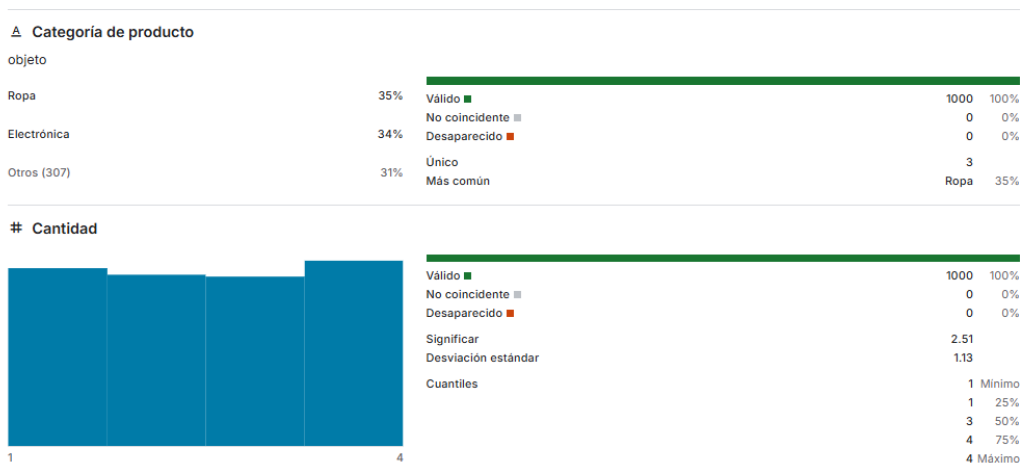
Los controles estadísticos de variables, la información recolectada sobre el comportamiento del proceso demanda vs abastecimiento muestra variación dentro de criterios definidos, al inicio del mes se aprecia variación, la cual puede deberse a causas no controlables. Lo mismo entre los periodos de última semana, lo que indica que el proceso puede estar saliendo de control en el límite superior, si comparamos los gráficos podemos mostrar que, si bien hay un cambio en el proceso, lo que indica una necesidad de revisión, no hay cambio repentino cada vez que se toma una muestra.



**Gráfico 4**

**Análisis de Desempeño**

Variables categoría de producto, la inclinación sigue siendo ropa con el mayor porcentaje de Participación entre los objetos mas consumidos, representa el 35% vs 34% en aparatos electrónicos.



## **Conclusiones y trabajos futuros**

El uso de Machine Learning sí mejora la toma de decisiones, especialmente en la toma la predicción del reabastecimiento automático. Se logro entrenar varios modelos capaces de predecir con buen nivel para analizar con detalles las demanda por categoría y el coste de estos, lo cual permitió entender cómo ciertas variables como categoría, costo, cantidad, preferencias del producto, juegan un rol Importante en el paretto de decisiones.

Finalmente, este trabajo permitió afianzar lo aprendido en el curso, desarrollar habilidades en análisis de datos, ahondar mas en el tema de Machine, lo cual es totalmente nuevo en el desarrollo profesional.

### **Referencias bibliográficas**

Herramientas de Mejoramiento Continuo para la Administración de la Calidad Carolina Cárdenas Vélez, Facultad de Ciencias Básicas e Ingeniería Corporación Universitaria Remington, Primera edición 2019

<https://www.kaggle.com/datasets/mohammadtalib786/retail-sales-dataset/data>

<http://www.scielo.org.co/pdf/rfiua/n43/n43a12.pdf>

<https://www.mecalux.com.co/blog/machine-learning-logistica>