



TRABAJO DE GRADO

Opción Seminario-Diplomado.

ALGORITMO COMPUTACIONAL PARA EL ANALISIS Y TOMA DE DECISIONES EN DATOS DE VOLUMEN EN VENTAS DE COMPAÑÍA DE PRODUCCION INDUSTRIAL, EMPLEANDO ESTRATEGIAS DE MACHINE LEARNING

Facultad de ingenierías

Especialización en dirección de operaciones y mejoramiento continuo

Estudiante: Jorge Hernán Londoño Trejos

Tutor: Juan Carlos Briñez de León

Opción de Trabajo de grado Seminario-Diplomado.

2025.

Contenido

Resumen 3

Palabras clave 4

Marco conceptual y contextual..... 5

Sector de balanceados 5

Que es Machine learning 5

¿Cómo funciona el Machine Learning?..... 6

Objetivos: 8

Objetivo general..... 8

Objetivos específicos 8

Desarrollo e implementación 9

Creando modelo de ML para la segmentación de clientes 10

Procesamiento de datos 22

Resultados: 22

Modelo de toma de decisiones 23

Análisis de datos 23

Conclusiones: 23

Bibliografía:..... 24

Resumen

El proyecto titulado " algoritmo computacional para el análisis y toma de decisiones en datos de volumen en ventas de una compañía de producción industrial, empleando estrategias de machine learning", se desarrolla con el objetivo de identificar, analizar y fortalecer conocimientos adquiridos para el análisis de datos computacionales, que son procesados a través de la configuración de algoritmos, que luego nos darán una análisis detallado por medio de histogramas o tablas, para la toma de decisiones.

A medida que avanza la lectura de este proyecto, se comparte un contexto general del mercado de balanceados a nivel nacional, datos aproximados del nivel de producción en el que se encuentra el sector, así mismo las oportunidades que con base en las cifras mencionados, amplía el panorama de la oportunidad de crecimiento en el que se encuentran las empresas que se dedican a la actividad elaboración y comercialización de alimento balanceado para animales.

El proyecto inicia con la consolidación de un grupo de datos, basado en las ventas de una compañía de elaboración de alimento balaceado del municipio de Bello; esta data contiene registros de 6 meses de ventas de las diferentes líneas de producto que normalmente se encuentran en la industria. Estos datos son procesados por medio de la asignación de algoritmos computacionales. Para ello se realiza una limpieza de la información, descartando variables que puedan afectar el procesamiento de los datos, transformando aquellos registros que se encuentren en formato alfanumérico a datos numéricos.

Se asignan los algoritmos computaciones para filtrar los datos.

Luego se aplica el modelo tipo Kmeans, para segmentar los datos según las variables, en este caso, se espera realizar los agrupamientos a tan solo 6 grupos.

Se procede a asignarles etiquetas (número del grupo) a la columna de línea de producto, buscando transformar el formato de datos de texto a numérico y con ello facilitar los demás pasos.

Se genera el histograma para tener una visual de los agrupamientos que realizó, basado en la serie de algoritmos que se asignaron. Se identifican 3 grupos representativos, frente a otros 3 que pueden tener un a oportunidad de mejora.

Partiendo del modelo de clustering aplicado, se asignan unas decisiones según los grupos previamente identificados, a los que se les otorga una serie de beneficios, y otros algunas condiciones específicas debido a las limitaciones a nivel económico y de comportamiento en la compra que pueda tener el cliente.

Al final se comparte una serie de resultados obtenidos y las pautas a tener en cuenta al momento de la generación de la orden de compra, según las especificaciones técnicas y de proceso interno del cliente final.

Palabras clave

Línea de producto, tabla de datos, Análisis de datos, Machine learning, Segmentación de datos, Clustering.

Marco conceptual y contextual

Sector de balanceados

Para el año 2024 los aportes de las compañías de balanceado registraron un crecimiento del 4,5% durante los últimos 6 años. La ANDI identificó que los principales aumentos se registraron para las líneas productivas de avicultura con un 65% y porcicultura con un 17%, seguidos de la ganadería y la línea de mascotas, con un 8% y 5% respectivamente.

Al analizar el comportamiento de la economía colombiana en el tercer trimestre de 2024, encontramos tasas positivas en los sectores de: actividades artísticas y de entretenimiento y actividades de los hogares (14,1); agricultura y ganadería (10,7%), actividades financieras y de seguros (4,4%), construcción (4,1%); administración pública y defensa (2,1%); actividades inmobiliarias (2,0%); comercio, transporte, alojamiento y servicios de comida (1,0%), electricidad, gas y agua (1,0%); información y comunicaciones (0,8%); y actividades profesionales y científicas (0,5%). (balance 2024 perspectivas 2025 - Asociación Nacional de Empresarios de Colombia ANDI)

Que es Machine learning

El machine learning, es definida como una disciplina del campo de la inteligencia artificial IA, definida a través de algoritmos que permite a los equipos computacionales identificar patrones de base de datos, analizar comportamientos y predicciones a partir de estos.

Este concepto a tenido un gran auge durante los últimos años, especialmente por la



necesidad de responder a cambios comportamentales y tecnológicos que se vienen presentando, facilitados por el boom informativo.

¿Cómo funciona el Machine Learning?

Hay cuatro etapas principales en el desarrollo de un modelo de Machine Learning. Por lo general, es un Data Scientist quien gestiona y supervisa el proceso.

El primer paso es seleccionar y preparar un conjunto de datos de entrenamiento. Esos datos se utilizarán para alimentar el modelo de Machine Learning para aprender a resolver el problema para el que se ha diseñado.

Los datos se pueden etiquetar para indicarle al modelo las características que debe identificar. También pueden estar sin etiquetar, entonces será el modelo el que deberá detectar y extraer características recurrentes por sí mismo.

En ambos casos, los datos deben prepararse, organizarse y limpiarse cuidadosamente. De lo contrario, el entrenamiento del modelo de Machine Learning puede estar sesgado. Los resultados de sus predicciones futuras se verán afectados directamente.

El segundo paso es seleccionar un algoritmo para ejecutar sobre el conjunto de datos de entrenamiento. El tipo de algoritmo que se emplea depende del tipo y del volumen de datos de entrenamiento y del tipo de problema que haya que resolver.

El tercer paso es entrenar el algoritmo. Es un proceso de repetición. Las variables se ejecutan a través del algoritmo y los resultados se comparan con los que debería haber producido. Los «pesos» y el sesgo se pueden ajustar para aumentar la precisión del resultado.

Después se vuelve a ejecutar las variables hasta que el algoritmo produzca el resultado correcto en la mayoría de los casos. El algoritmo entrenado es el modelo de Machine Learning.

El cuarto y último paso es el uso y la mejora del modelo. Utilizamos el modelo sobre nuevos datos, cuyo origen depende del problema que haya que resolver. Por ejemplo, en los correos electrónicos se usará un modelo de Machine Learning diseñado para detectar spam. (datascientest, 2024)

Objetivos:

Objetivo general

Implementar un algoritmo computacional para el análisis y toma de decisiones en datos de volumen de ventas de compañía de producción industrial, empleando estrategias de machine learning

Objetivos específicos

Consolidar los datos extraídos del histórico de ventas generadas durante el segundo semestre del año 2023 para una compañía de alimento balanceado de la ciudad de Bello - Antioquia.

Implementar un algoritmo de Machine Learning para la toma de decisiones a partir de datos suministrados por los clientes a la hora de programar un pedido.

Asignar las condiciones de pago del cliente a partir del comportamiento de compra y volumen según orden de compra.

Analizar el comportamiento de los algoritmos definidos para la toma de decisiones.

Desarrollo e implementación

Se extrae la base de datos suministrada por una compañía de alimento balanceado con sede principal en el municipio de Bello – Antioquia, relacionado a las ventas registradas durante el segundo semestre del año 2023, se determinan las siguientes variables para su análisis respectivo:

- **Mes:** fecha definida en número para identificar la fecha de generación de la factura de venta.
- **C.O:** asociado a la ciudad de destino del pedido solicitado por el cliente.
- **Bodega:** sede de la compañía por la que el cliente retira su pedido (contiene una cobertura nacional)
- **Semana:** corresponde a la fecha estimada durante el último semestre para la entrega de la orden de compra.
- **Referencia:** tipo de producto solicitado en la orden de compra
- **Cantidad:** las unidades solicitadas por referencia que se tienen en cuenta para la programación de la producción.
- **Valor neto:** el valor en \$ de la orden de compra
- **Línea de producto:** corresponde a la segmentación del mix de productos suministrado por la compañía.

Se emplea la técnica de clustering (aprendizaje no supervisado) cuyo objetivo es identificar patrones no identificados en los datos. El método agrupa y relaciona los registros según similitudes en conjuntos llamados clústeres; lo anterior permite analizar la estructura de la información y segmentar los datos según los patrones existentes.

Creando modelo de ML para la segmentación de clientes

Se consolidaron los registros, relacionados con el comportamiento de ventas registradas durante el segundo semestre del año 2023, buscando agrupar los clientes según las variables anteriormente mencionadas, a través del método del codo (Elbow method).

Algoritmo empleado:

```
import matplotlib.pyplot as plt
import numpy as np
import time
from sklearn.cluster import KMeans
from sklearn import metrics

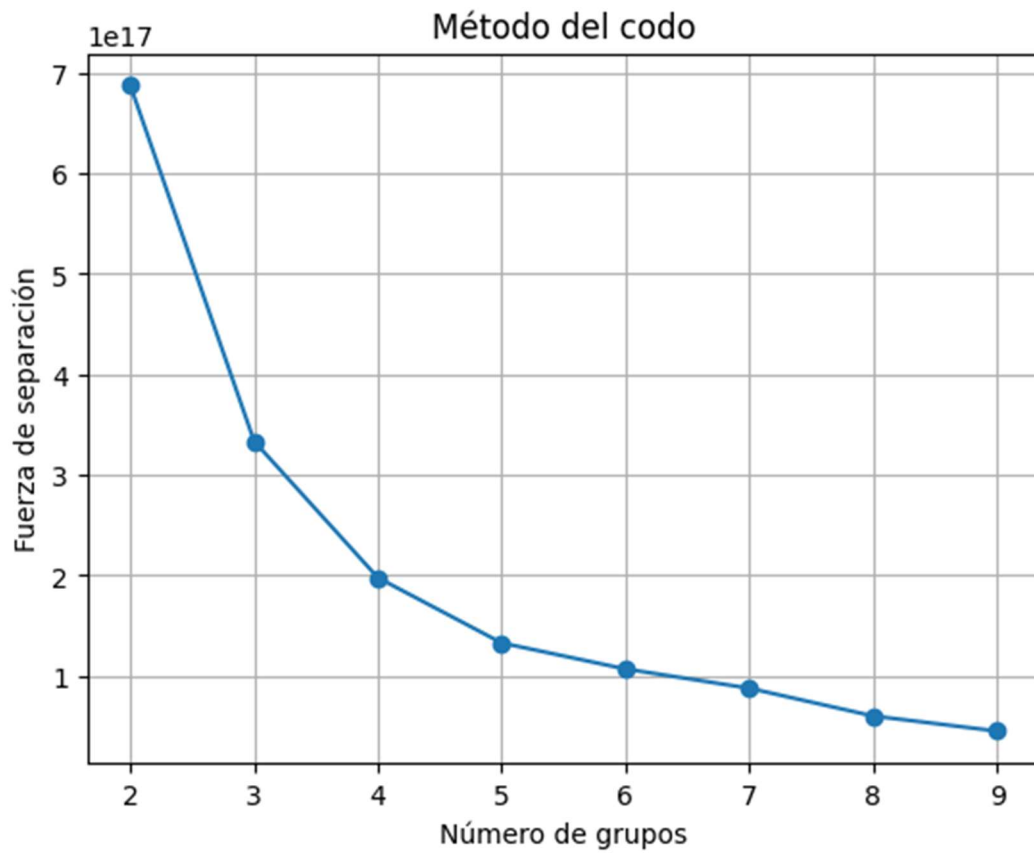
# Ejecutar KMeans con diferentes números de clusters
inertia_values = []
silhouette_scores = []
Inicial = 2
Final = 10
k_range = range(Inicial, Final)

for k in k_range:
    kmeans = KMeans(n_clusters=k, n_init = 'auto', random_state=42)
    kmeans.fit(Datos_Array)
    inertia_values.append(kmeans.inertia_)
    silhouette_scores.append(metrics.silhouette_score(Datos_Array, kmeans.labels_))

# Graficar el método del codo
plt.plot(k_range, inertia_values, marker='o')
plt.xlabel('Número de grupos')
plt.ylabel('Fuerza de separación')
plt.title('Método del codo')
plt.grid(True)
plt.show()

print('')
```

1 Imagen



2 Imagen

En la imagen 2 se definen los cluster asignados, proyectados a 9 grupos iniciales según la línea de producto tales como: 'porcicultura', 'avicultura', 'engorde', 'piscicultura', 'mascotas', 'otras especies menores', 'ganadería', 'sales', 'equinos', 'mascotas alimento', 'mascotas arena' durante el segundo semestre del año 2023.

El código empleado fue realizado en una herramienta de Google colab.

Se crea el modelo de grupos a partir del algoritmo:

```
from sklearn.cluster import KMeans
k = int(input('Ingrese el número de grupos deseados: '))
Modelo_Cluster = KMeans(k, random_state=37)
Modelo_Cluster.fit(Datos_Array)
```

2v1Imagen

Se solicita al algoritmo distribuir los datos en 6 grupos:

```
Ingrese el número de grupos deseados: 6
KMeans
KMeans(n_clusters=6, random_state=37)
```

3 Imagen

Se procede a asignar una nueva etiqueta a través de los algoritmos que permita crear una columna que asigne el grupo según las condiciones dadas:

```
Etiquetas = Modelo_Cluster.predict(Datos_Array)
Mis_datos2=Mis_datos.copy()
Mis_datos2['Grupo'] = Etiquetas
Mis_datos2.head(10)
```

31Imagen

Así mismo la solicitud de 10 líneas de registros que nos permitan graficar a través de una tabla los datos anteriormente establecidos:

	MES	c.o	Bodega	semana	Referencia	Cantidad	Valor neto	LINEA PRODUCTO	Grupo
0	7	38	67	27	31563	250	0	1	5
1	7	40	67	27	30403	20	0	1	5
2	7	40	67	27	30408	25	0	1	5
3	7	40	67	27	30745	30	0	1	5
4	7	40	67	27	31558	25	0	1	5
5	7	40	67	28	30408	40	0	1	5
6	7	40	67	28	30403	20	0	1	5
7	7	40	67	28	31559	15	0	1	5
8	7	40	67	28	31558	25	0	1	5
9	7	40	67	29	30408	12	0	1	5

4 Imagen

La distribución de los grupos genera un conteo aproximado según condiciones mínimas y máximas de los datos recolectados.

	MES	c.o	Bodega	semana	Referencia	Cantidad	Valor neto	LINEA PRODUCTO	Grupo
count	48985.000000	48985.000000	48985.000000	48985.000000	48985.000000	48985.000000	4.898500e+04	48985.000000	48985.000000
mean	9.002797	3.328897	37.882393	37.320874	40191.334694	520.240604	3.962481e+06	3.954986	3.740737
std	1.997061	7.981401	32.593391	8.808859	17379.257128	2366.109712	7.026946e+06	4.191529	1.944746
min	6.000000	1.000000	1.000000	22.000000	10204.000000	1.000000	0.000000e+00	1.000000	0.000000
25%	7.000000	1.000000	1.000000	30.000000	31561.000000	5.000000	2.330690e+05	1.000000	3.000000
50%	9.000000	1.000000	67.000000	37.000000	32207.000000	17.000000	1.044660e+06	1.000000	5.000000
75%	11.000000	1.000000	67.000000	45.000000	50107.000000	60.000000	4.260207e+06	9.000000	5.000000
max	12.000000	47.000000	67.000000	52.000000	91576.000000	18630.000000	1.175328e+08	12.000000	5.000000

5 Imagen

La tabla anterior nos muestra que para las ordenes de compra en las que hay un valor de venta cercano a \$0, el algoritmo lo clasifica en el grupo 0 (mínimo); este grupo

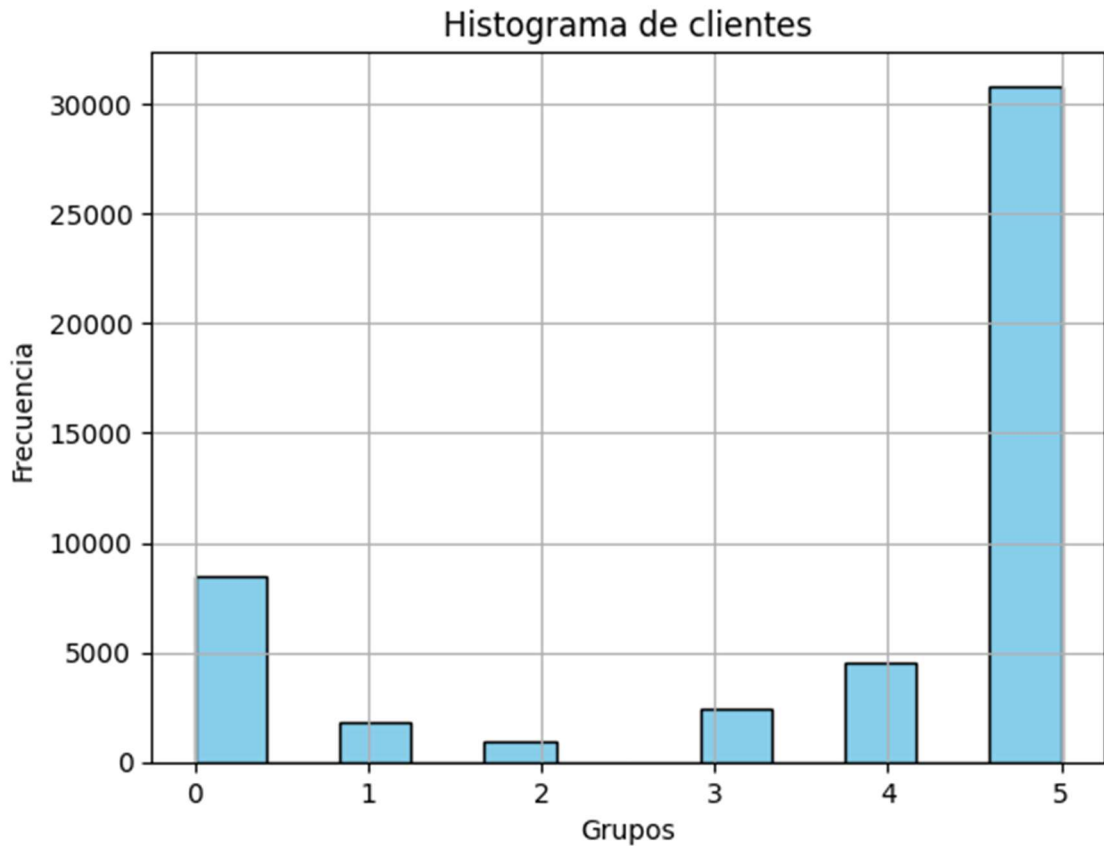
está asociado a los clientes que hacen parte del holding corporativo y que de acuerdo a las condiciones comerciales tiene un registro en la factura de cobro \$0. Para el caso de los máximos, segmenta los datos según los clientes que registren solicitudes cercanas a las 18 mil unidades.

Para llevar los datos anteriores a una visual mas clara, se emplea el siguiente algoritmo que permite generar un histograma de clientes, clasificando los grupos vs las unidades solicitadas (frecuencia):

```
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd

# Histograma de la variable 'Grupo'
plt.hist(Mis_datos2['Grupo'], bins=12, color='skyblue', edgecolor='black')
plt.xlabel('Grupos')
plt.ylabel('Frecuencia')
plt.title('Histograma de clientes')
plt.grid(True)
plt.show()
```

6|Imagen



71Imagen

El histograma anterior nos muestra que el grueso de los clientes está segmentado en el grupo 5, y los de menor presencia en el registro de venta están clasificados en los grupos 1 y 2.

Continuamos con la generación del punto de equilibrio o promedio de los datos con el siguiente algoritmo:

```
# Obtener los centroides del modelo K-Means
Centroides=Modelo_Cluster.cluster_centers_

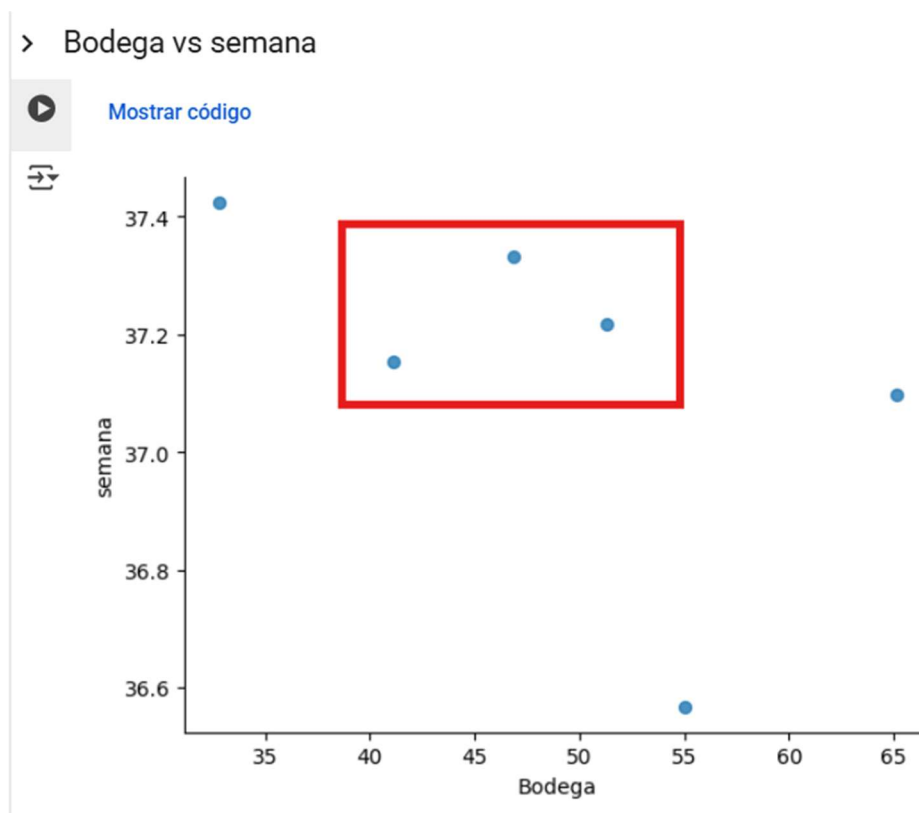
# Obtener los nombres de las columnas del dataset original
nombres_variiables = Mis_datos.columns[:] # Excluir la última columna que contiene los clusters

# Crear un DataFrame para los centroides con los nombres de las columnas
Centroides_df = pd.DataFrame(Centroides[:,:] , columns=nombres_variiables)

# Imprimir los centroides con los nombres de las variables
print("Centroides:")
Centroides_df.head(k)
```

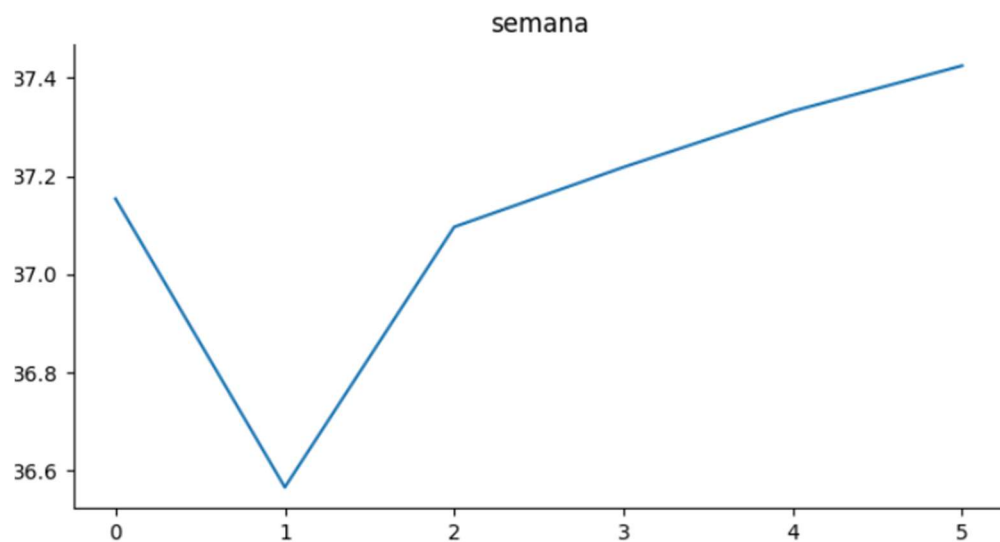
8Imagen 9

Tomamos como ejemplo el grafico Bodega vs Semana, que nos permite identificar las bodegas que mayor interacción presentaron durante las semanas 36 y 37:



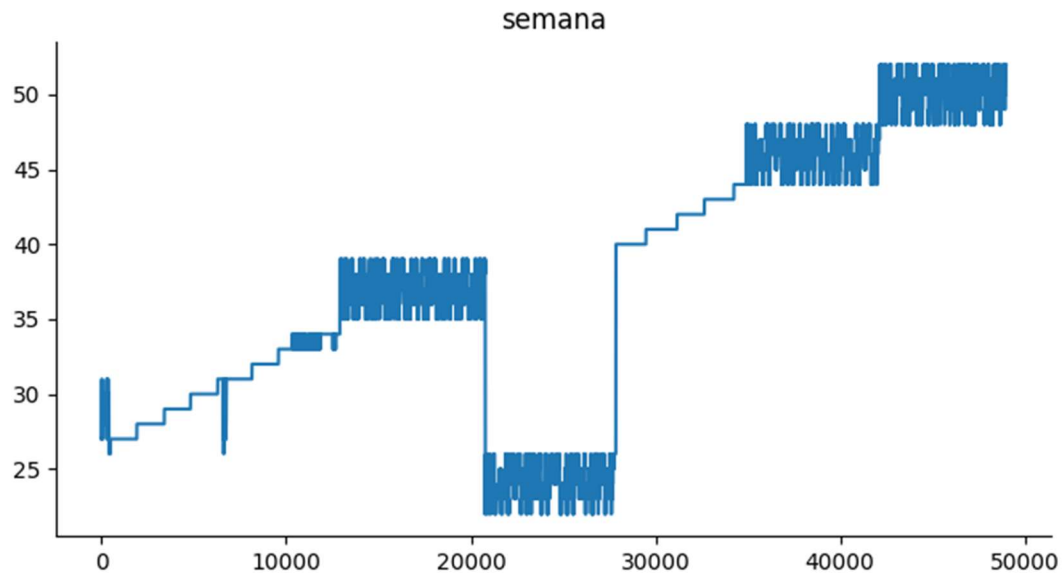
9Imagen 10

En el siguiente histograma vemos que el grupo 1 no registra venta al final de la semana 36 y el resto de la semana 37, contrario a los grupos 4 y 5 que presentan un comportamiento de compra importante en este periodo:



Asi mismo se identifico la concentración de cantidades en las órdenes de compra, para ello se emplea el algoritmo que me permita hallar este comportamiento:

```
from matplotlib import pyplot as plt
Mis_datos2['semana'].plot(kind='line', figsize=(8, 4), title='semana')
plt.gca().spines[['top', 'right']].set_visible(False)
```



En nuestro modelo de clustering empleamos los datos de nuestro dataset y los adaptamos de manera que pudieran ser analizados de forma numérica, bajo las variables de valor de la orden de compra, cantidades programadas y la estimación de fecha de retiro de la solicitud. Con estos datos logramos definir unos patrones que de pueden identificar y segmentar los clientes según los grupos establecidos anteriormente, y así tomar las acciones necesarias para el proceso interno de la empresa, a nivel de producción, logística y recuperación de cartera:

La segmentación definida por los algoritmos empleamos nos permiten definir políticas de inventarios, programación de pedidos y de recuperación de cartera.

Al modelo se le asignan 8 variables:

1. Ingrese el mes de reabastecimiento: 7: jul - 8: agos - 9: sept - 10: oct
- 11: nov - 12: dic

2. Ingrese el # de ciudad destino:

1: Medellin

4: Cartago

6: Bello

21: San pedro

2: Don Matías

27: La Pintada

28: Argelia

30: Ebéjico

36: Manizales

37: Pereira

38: Calarcá

40: Montería

44: Barranquilla

46: Cartagena

47: Cordoba

3. Digite la Bodega en que recoge pedido:

1: Bello

45: Cartago

67: Medellin

4. Digite la semana en que retira su pedido (rango de 22 - 52)
5. Digite la referencia a solicitar (desde 10204 hasta 91576)
6. Digite cantidad a solicitar desde 1 hasta 18630 unidades
7. Digite el abono realizado, valor en miles de \$0 hasta \$100 millones
8. Ingrese la línea que desea comprar:

1 PORCICULTURA

2 AVICULTURA

9 ENGORDE

6 PISCICULTURA

10 MASCOTAS - MASCOTAS ALIMENTO - MASCOTAS ARENA

11 OTRAS ESPECIES MENORES

12 GANADERIA

4 SALES

7 EQUINOS

Algoritmo propuesto:

```

Cliente_New = np.zeros((1,8)) #un cliente con 8 variables
Cliente_New[0,0] = float(input('Ingrese el mes de reabastecimiento: 7:jul - 8:agos - 9:sept - 10:oct - 11:nov - 12:dici'))
Cliente_New[0,1] = float(input('Ingrese el # de ciudad destino: 1:Medellin 4:Cartago 6:Bello 21:San pedro 2:Don Matias 27:La Pintada 28:Argelia 30:Ebejico 36:Manizales 37:Pereira 38:Calarcá 40:Monteria 44:Barranquilla 46:Cartagena 47:Cordoba'))
Cliente_New[0,2] = float(input('Digite la Bodega en que recoge pedido: 1:Bello 45:Cartago 67:Medellin: '))
Cliente_New[0,3] = float(input('Digite la semana en que retira su pedido (rango de 22 - 52):'))
Cliente_New[0,4] = float(input('Digite la referencia a solicitar (desde 10204 hasta 91576):'))
Cliente_New[0,5] = float(input('Digite cantidad a solicitar desde 1 hasta 18630 unidades:'))
Cliente_New[0,6] = float(input('Digite el abono realizado, valor en miles de $0 hasta $100 millones:'))
Cliente_New[0,7] = float(input('Ingrese la línea que desea comprar: 1 PORCICULTURA - 2 AVICULTURA - 3 ENGORDE - 6 PISCICULTURA - 10 MASCOTAS - 11 OTRAS ESPECIES MENORES - 12 GANADERIA - 4 SALES - 7 EQUINOS - 10 MASCOTAS ALIMENTO - 10 MASCOTAS ARENA 10'))

Etiqueta_cliente = Modelo_cluster.predict(Cliente_New)

print('Según los datos del cliente, el grupo es: ',Etiqueta_cliente)
print(' ')

if Etiqueta_cliente == 0:
    print('cliente es estable y cumplido con fecha de retiro del pedido, se debe anticipar producción con 2 días según promesa de retiro, dar crédito de 15 días:')
if Etiqueta_cliente == 1:
    print('Storage pago de contado, con descuento del 3% pie factura')
if Etiqueta_cliente == 2:
    print('cliente con potencial crecimiento, hacer acompañamiento técnico semanal, descuento del 2%')
if Etiqueta_cliente == 3:
    print('Sin descuento, condición de pago de contado')
if Etiqueta_cliente == 4:
    print('Programar producción, asignar 10 días de crédito y descuento del 3% en pie factura')
if Etiqueta_cliente == 5:
    print('Cuenta en participación, no requiere descuento, Programar producción urgente hace parte del Holding:')

```

Ingrese el mes de reabastecimiento: 7:jul - 8:agos - 9:sept - 10:oct - 11:nov - 12:dici8
Ingrese el # de ciudad destino: 1:Medellin 4:Cartago 6:Bello 21:San pedro 2:Don Matias 27:La Pintada 28:Argelia 30:Ebejico 36:Manizales 37:Pereira 38:Calarcá 40:Monteria 44:Barranquilla 46:Cartagena 47:Cordoba1
Digite la Bodega en que recoge pedido: 1:Bello 45:Cartago 67:Medellin: 1
Digite la semana en que retira su pedido (rango de 22 - 52):29
Digite la referencia a solicitar (desde 10204 hasta 91576):40513
Digite cantidad a solicitar desde 1 hasta 18630 unidades:15000
Digite el abono realizado, valor en miles de \$0 hasta \$100 millones:13000000
Ingrese la linea que desea comprar: 1 PORCICULTURA - 2 AVICULTURA - 9 ENGORDE - 6 PISCICULTURA - 10 MASCOTAS - 11 OTRAS ESPECIES MENORES - 12 GANADERIA - 4 SALES - 7 EQUINOS - 10 MASCOTAS ALIMENTO - 10 MASCOTAS ARENA 102
Según los datos del cliente, el grupo es: [3]

Sin descuento, condicion de pago de contado

En la gráfica, podemos ver las condiciones que se le dieron a cada variable:

Cliente reabastece en el mes 8 - Agosto

de la ciudad de destino es: 1 Medellín

Bodega en que recoge pedido: 1 Bello

Semana en que retira su pedido (rango de 22 - 52): 29

Referencia por solicitar (desde 10204 hasta 91576): 40513

Cantidad por solicitar desde 1 hasta 18630 unidades:15.000 unds

Digite el abono realizado, valor en miles de \$0 hasta \$100 millones: \$ 13.000.000

de Linea que desea comprar: 2 AVICULTURA

La decisión que nos arroja el modelo es que:

Según los datos del cliente, el grupo es: [3]

Es decir, con las condiciones que se le asignaron en modelo de clustering, el cliente cumple las condiciones de la variable 3: **Sin descuento, condición de pago de contado**

Analizando a profundidad, este cliente tiene un comportamiento de compra que moderada si se tiene en cuenta el abono realizado por 13 millones, una cifra

Procesamiento de datos

La base de datos se trasladó a un formato Excel, se carga la información a la herramienta Google colab, y se organizaron mediante correcciones de errores, se suprime columnas indeseadas tales como aquellas con formato fecha, porque ya se tiene en cuenta el mes; se identifican variables en formato texto, especialmente de línea de producto y se reemplazan dichos valores en números.

Resultados:

Mes de reabastecimiento	# ciudad destino	Bodega	Semana	ref.	cant.	abono	línea producto	Decisión
9	36	45	40	30303	8.050	\$ 66.000.000	10	cliente con potencial crecimiento, hacer acompañamiento tecnico semanal, descuento del 2%
8	47	67	46	31561	4.000	\$ 89.000.000	3	cliente con potencial crecimiento, hacer acompañamiento tecnico semanal, descuento del 2%
7	2	1	39	40513	3.900	\$ 42.600.000	4	cliente con potencial crecimiento, hacer acompañamiento tecnico semanal, descuento del 2%
12	4	2	49	31563	800	\$ 8.000.000	3	Programar producción, asignar 10 dias de credito y descuento del 3% en pie factura
8	46	1	27	32207	250	\$ 19.000.000	3	Otorgar pago de contado, con descuento del 3% pie factura

Se aplica el modelo a 5 clientes que generan ordenes de compra para las próximas semanas, en diferentes regiones del país. En la gráfica, se identifica que los clientes que son mas moderados en los pedidos y a su vez registran abonos, se enlistan en fila preferencial para programar la producción de sus pedidos, y con ello se puede obtener un recaudo de la cartera mucho más rápido y mejorar el indicado de rotación de producto en almacén.

Para los primeros clientes que registraron pedidos de mas de 3.900 unidades (primeras 3 lineas), son clientes que para la fecha de generado este prototipo de predicción, están proyectados a un crecimiento importante durante las próximas semanas; especialmente porque las ordenes de compra que ingresan, no tienen fecha de entrega de manera inmediata, es decir el cliente tiene estimado pasar pedido durante las siguientes fechas según el mes de reabastecimiento y a su vez notifican a la compañía

con una fecha puntual (semana) de retiro del pedido. Lo anterior alinea las proyecciones y la programación de materiales en la planta de producción, es decir, toda la logística de suministro, para lograr cumplir con el pedido según la promesa de retiro de la mercancía por parte del cliente.

Modelo de toma de decisiones

Análisis de datos

El análisis de los datos se realizó mediante la implementación de un modelo de clustering, buscando segmentar los datos según unos patrones no identificados, así mismo determinar las acciones a emplear para atender los requerimientos de clientes potenciales, basados en la orden de compra enviada, con estos datos, nos permite identificar los pasos a seguir, en el arco de la programación de ordenes de producción, agendamiento de despachos y gestión de cartera con el cliente. Facilitando el perfil del cliente y los beneficios que se le pueden otorgar, según las condiciones asignadas.

Los gráficos y estadísticas nos permiten identificar comportamientos y agrupar los clientes a partir del histórico de ventas. Las solicitudes que lleguen a la compañía son una muestra del comportamiento de la facturación que se puede dar para los próximos 6 meses.

Conclusiones:

Un modelo de predicción, fortalece los procesos internos de la compañía, desde el ámbito industrial, control de calidad y gestión logística.

Se mejora el nivel de servicio, lo que nos permite la disponibilidad del

producto. Con esto se estima que la producción este acentuada a la realidad según el comportamiento de venta.

Las variables seleccionadas nos permiten ser precisos con la solicitud del cliente. El servicio de clustering es principalmente de acompañamiento, para la toma de decisiones puntuales.

La segmentación de los datos nos permite identificar el volumen mínimo de inventario que se debe contar para cumplir con los requerimientos de cliente.

La toma de decisiones se basa principalmente con la información suministrada por el cliente, y según a los datos entregados por el cliente, el algoritmo ayuda a tomar esa decisión de manera inmediata, sin necesidad de esperar respuestas con otros procesos.

Bibliografía:

- https://porkcolombia.co/indicadores_sector/boletin-economico-enero-noviembre-2024/
- <https://www.andi.com.co/home/camara/17-industria-de-alimentos-balanceados>
- Balance 2024 perspectivas 2025 - Asociación Nacional de Empresarios de Colombia ANDI
- Hidalgo, Andrés, Comenzando con Numpy y Pandas, junio 2024.