

2015-06-01

Sistema de información para administración de inventarios basado en técnicas de minería de datos

Diana Lancheros-Cuesta

Universidad de La Salle, Bogotá, dilancheros@unisalle.edu.co

Andrés David Mojica Ospina

Universidad de La Salle, Bogotá, andavidosp@gmail.com

Daniel Alberto Barbosa Vargas

Universidad de La Salle, Bogotá, dabarva@hotmail.com

Follow this and additional works at: <https://ciencia.lasalle.edu.co/ep>

Citación recomendada

Lancheros-Cuesta, Diana; Mojica Ospina, Andrés David; and Barbosa Vargas, Daniel Alberto (2015)
"Sistema de información para administración de inventarios basado en técnicas de minería de datos,"
Épsilon: Iss. 24 , Article 7.

Disponible en:

This Artículos de investigación is brought to you for free and open access by the Revistas descontinuas at Ciencia Unisalle. It has been accepted for inclusion in Épsilon by an authorized editor of Ciencia Unisalle. For more information, please contact ciencia@lasalle.edu.co.

Sistema de información para administración de inventarios basado en técnicas de minería de datos

DIANA LANCHEROS-CUESTA¹

ANDRÉS DAVID MOJICA OSPINA²

DANIEL ALBERTO BARBOSA VARGAS³

RESUMEN

Actualmente existen diferentes sistemas de información que, a partir de grandes grupos de datos, pueden predecir comportamientos y permiten la toma de decisiones en las empresas; sin embargo, son herramientas que resultan difíciles de adquirir por su complejidad y gran robustez, debido a los costos elevados o a las restricciones legales. En las organizaciones, la administración de la información se convierte en una actividad prioritaria, por cuanto una buena gestión en la manipulación de los datos mejora los procesos. En línea con lo anterior, este artículo muestra el desarrollo de un *software* basado en técnicas de minería de datos (*datamining*) que permite la recopilación y el almacenamiento de información, con el fin de analizar patrones ocultos de comportamiento en el departamento de inventarios de una empresa.

Palabras clave: minería de datos, lógica difusa, inventario, sistema de información.

¹ Doctora en Ingeniería por la Pontificia Universidad Javeriana, Bogotá, Colombia. Magíster en Tecnologías de la Información por la Universidad Pedagógica Nacional, Bogotá, Colombia. Ingeniera de Diseño y Automatización Electrónica por la Universidad de La Salle, Bogotá, Colombia. Actualmente se desempeña como docente en el programa de Ingeniería en Automatización de la Universidad de La Salle, Bogotá, Colombia. Correo electrónico: dilancheros@unisalle.edu.co

² Ingeniero en Automatización por la Universidad de La Salle, Bogotá, Colombia. Correo electrónico: andavidosp@gmail.com

³ Ingeniero en Automatización por la Universidad de La Salle, Bogotá, Colombia. Correo electrónico: dabarva@hotmail.com

FECHA DE RECEPCIÓN: 6 DE ABRIL DEL 2014 • FECHA DE ACEPTACIÓN: 5 DE OCTUBRE DEL 2014

Cómo citar este artículo: Lancheros-Cuesta, D., Mojica Ospina, A. D. y Barbosa Vargas, D. A. (2015). Sistema de información para administración de inventarios basado en técnicas de minería de datos. *Épsilon*, 24, 143-173.

Information System for Inventory Management Based on Data Mining Techniques

ABSTRACT

There are currently different information systems that can predict behavior from large sets of data and which allow decision-making in companies; however, these tools are difficult to acquire due to their complexity and robustness, as well as high costs or legal restrictions. Information management in companies becomes a priority activity, because good management in data handling improves processes. Accordingly, this paper shows the development of a software based on data mining techniques, thus enabling the collection and storage of information in order to analyze hidden behavior patterns in a company's inventory department.

Keywords: data mining, fuzzy logic, inventory, information system.

Sistema de informação para administração de inventários baseado em técnicas de mineração de dados

RESUMO

Atualmente existem diferentes sistemas de informação que, a partir de grandes grupos de dados, podem prever comportamentos e permitem a tomada de decisões nas empresas; contudo, são ferramentas que resultam difíceis de adquirir por sua complexidade e grande robustez, devido aos custos elevados ou às restrições legais. Nas organizações, a administração da informação torna-se uma atividade prioritária, por isso, uma boa gestão na manipulação dos dados melhora os processos. Na mesma linha do anterior, este artigo mostra o desenvolvimento de um *software* baseado em técnicas de mineração de dados (*datamining*), que permite a recopilación e o armazenamento de informação, com a finalidade de analisar padrões ocultos de comportamento no departamento de inventários de uma empresa.

Palavras chave: mineração de dados, lógica difusa, inventário, sistema de informação.

Introducción

El desarrollo de un sistema de información eficiente que precise la extracción de datos de una empresa y que represente mayor economía hacia inversionistas constituye la pauta primordial para llevar a cabo la implementación de sistemas de información en la industria actual.

Actualmente, la necesidad de utilizar técnicas de minería de datos mediante un *software* determinado obliga a los ingenieros y desarrolladores a revalidar los conocimientos adquiridos y aplicados al diseño de sistemas de información. Por tanto, es necesario indicar los trabajos relacionados con este tipo de sistemas, como se reseña a continuación.

Moreno (2005) desarrolló un sistema de información modular para la compra, la venta y el inventario de una fábrica de calzado. En este trabajo se muestra la forma de implementar una terminal móvil para el control de inventarios, así como el diseño de las bases de datos, en conjunto con el sistema de información, para integrar, al final, todos los componentes del proyecto. Aquí se tienen en cuenta básicamente las necesidades transaccionales de un sistema de información.

Luna *et al.* (2009) implementaron un sistema de planificación de recursos empresariales (ERP) para optimizar sistemas de gerencia de restaurantes o bares protegidos por una llave electrónica. El desarrollo consistió en el reagrupamiento de todos los datos con el objetivo de mejorar la gestión de los datos. En primera instancia, los autores determinaron los requerimientos funcionales y los casos de uso del sistema. El trabajo evidencia la necesidad de establecer puntos de información de la base de datos, donde la información tenga nivel de seguridad gestionada por el administrador.

Por su parte, Bozo (2009) desarrolló un sistema de gestión de inventarios, ventas y facturación para una empresa automotriz, que incluía un sistema de seguridad de acceso a los datos tanto en el sistema operativo como en el programa. Esto permitió definir y asignar los diferentes accesos a los usuarios de acuerdo con su función.

Vallejos (2006) realizó un estudio con la definición, las ventajas y las desventajas de las principales técnicas de minería de datos. De igual forma, Tapia (2006) determinó una metodología para sectorizar pacientes que consumen medicamentos,

y para ello aplicó *datamart* y *datamining* en un hospital nacional. La tabla 1 muestra una descripción de otros trabajos relacionados con minería de datos, así como otras metodologías en diferentes campos de aplicación y el porcentaje de efectividad que tuvo la técnica de minería de datos aplicada.

Tabla 1. Síntesis de metodologías de trabajos de grado relacionados con minería de datos

TÍTULO, AUTOR, AÑO	METODOLOGÍA	RESUMEN	EFFECTIVIDAD (%)
"Método general de extracción de información basado en el uso de lógica borrosa. Aplicación en portales web" (Roper, 2009)	Lógica fuzzy	En esta tesis se ha planteado un nuevo método de extracción de información basado en el uso de la lógica borrosa para un conjunto de conocimiento general. Este método se aplica, en particular, a portales web.	90
"Diseño e implementación de un sistema ERP para optimizar sistemas de gerencia de restaurantes o bares protegido por una llave electrónica" (Beltrán <i>et al.</i> , 2009).	Regresión lineal, árboles de decisión	En el sistema ERP se reagrupan todos los datos de la empresa. En el caso de la empresa seleccionada para el caso de estudio, este sistema hace que sean más dinámicas las funciones administrativas de la empresa. Con base en estas consideraciones, se establecieron los requerimientos funcionales y los casos de uso del sistema, para el manejo operativo de este cuando se elaboró la base de datos. Fue de vital importancia establecer puntos de información de la base de datos, donde esta tenga cierto nivel de seguridad por el administrador.	90
"Desarrollo metodológico y técnico de un pronóstico de ventas para un producto con patrones cíclicos" (Gracia, 2002)	Regresión lineal	En el documento se ha sugerido la posibilidad de utilizar modelos matemáticos edificados a través del comportamiento de las ventas. Estos son correlacionados con indicadores del comportamiento económico del país, y se edifica luego la ecuación de pronósticos con la ayuda de la herramienta estadística de análisis de regresión lineal.	90
"Diseño e implementación de un sistema de información modular para la compra, venta e inventario de una fábrica de calzado" (Moreno, 2005)	Modelos estadísticos	En el documento se muestra la forma de implementar al terminal móvil para el control de inventarios, así como el diseño de las bases de datos en conjunto con el sistema de información, para al final integrar todos los componentes del proyecto de forma adecuada.	90

TÍTULO, AUTOR, AÑO	METODOLOGÍA	RESUMEN	EFFECTIVIDAD (%)
"Formulación de minería de datos para la Empresa Distribuidora de Productos Espinoza Aguilar S.A." (Alcántara, 2012)	Modelo predictivo	Con el apoyo de la metodología de la minería de datos, se buscan tendencias que dan una mejor rentabilidad para el negocio. Con ello se reducen los costos logísticos mediante el uso de la información que se tiene almacenada históricamente del propio entorno del negocio.	80
"Desarrollo de un <i>software</i> para el control de inventario de productos terminados para los departamentos de atención al cliente, la línea de producción 'sector beta', y despacho en una empresa alimentos" (García, 2009)	Árboles de decisión	Consiste en el desarrollo de un <i>software</i> para la automatización de las operaciones de realización, aprobación y despacho de las ordenes de carga, así como la creación y actualización del inventario en la empresa alimentos Súper S planta Barcelona C.A. Este <i>software</i> facilitará los procesos llevados a cabo los operadores de los diferentes departamentos involucrados. Este proyecto plantea optimizar y reducir los tiempos y los errores en las operaciones llevadas a cabo en los departamentos de Atención al Cliente, Cuentas por Cobrar, Despacho, Línea de Producción, Romana, para lo cual se tienen en consideración los adelantos tecnológicos en el área de <i>software</i> .	70
"Aplicación de minería de datos para la exploración y detección de patrones delictivos en Argentina" (Perversi, 2007)	Árboles de decisión; algoritmo <i>ID3 K-means; clustering</i> ; algoritmo de inducción C4.5 para identificar reglas de pertenencia a cada uno de los grupos	En este trabajo se ha demostrado no solo que es factible aplicar minería de datos a la información criminal en Argentina, sino que también evidencia su alto valor agregado para el análisis y la generación de nuevo conocimiento.	65 (aún en estudio)

Fuente: elaboración propia.

Como se observa en la tabla 1, la efectividad varía según la metodología usada en cada una de las diferentes áreas. Por tal motivo, puede deducirse que la aplicación de la minería de datos en un ámbito de *software* para el apoyo en la toma de decisiones, con base en la programación lineal y la computación gráfica para la visualización de los datos, permite obtener resultados satisfactorios. A continuación se describe la metodología utilizada en el desarrollo del *software*.

Metodología

Para el desarrollo del *software* se llevaron a cabo las siguientes etapas: a) definición del problema, b) preparación los datos, c) exploración los datos, d) generación de modelos, e) exploración y validación de modelos, f) implementación y g) actualización del modelo final, como se muestra a continuación.

Definición del problema

Para el diseño y desarrollo del sistema se tuvo en cuenta la existencia de una serie de datos sobre las ventas en una empresa de servicios, los cuales están desorganizados. Adicionalmente, el sistema de información actual no permite que se modifiquen los datos de los productos de forma rápida y sin tener conocimientos de programación, lo cual genera reportes erróneos de la empresa y no permite la toma de decisiones.

Preparación y exploración de los datos

Los datos proporcionados por la empresa están dispersos y almacenados en un mismo archivo y presentan una cantidad de variables irrelevantes para el proceso. También contienen incoherencias, como entradas que faltan o no separación visual de la información. La figura 1 muestra una fracción del archivo de información de ventas, donde se nota la dificultad para utilizar estos datos de forma tal que el sistema brinde análisis claros y concisos. Por tal razón, es necesario ordenar y preparar nuevamente.

Generación de los modelos

Para la generación del modelo computacional que soporta el sistema de información se implementaron técnicas de control de inventarios mediante la estadística y lógica *fuzzy*.

Modelo basado en la estadística y el control de inventarios

Con el objetivo de implementar un sistema que provea al usuario la capacidad de tomar decisiones con base en métodos de administración de inventarios, y que a la vez permita realizar funciones diferentes del sistema, es necesario remitirse a

0553	11-1-2013	4987	SOKA MORA BOTELLA LITRO	
0332	12-1-2013	4987	SOKA MORA BOTELLA LITRO	
0223	13-1-2013	4987	SOKA MORA BOTELLA LITRO	
0333	14-1-2013	4987	SOKA MORA BOTELLA LITRO	
0444	15-1-2013	4987	SOKA MORA BOTELLA LITRO	
0555	16-1-2013	4987	SOKA MORA BOTELLA LITRO	
0666	17-1-2013	4987	SOKA MORA BOTELLA LITRO	
0777	18-1-2013	4987	SOKA MORA BOTELLA LITRO	
0888	19-1-2013	4987	SOKA MORA BOTELLA LITRO	
0999	20-1-2013	4987	SOKA MORA BOTELLA LITRO	
1234	16-2-2013	1760	NECTAR BABY MANZANA 124G	
1234	16-2-2013	2400	FRUTTO CLASICO PERA CAJA 200ML	
0022	24-2-2013	1760	NECTAR BABY MANZANA 124G	
0022	24-2-2013	2351	FRUTTO CLASICO PERA CAJA LITRO	
0022	24-2-2013	2401	FRUTTO CLASICO MANZANA CAJ 200ML	
0015	24-2-2013	1760	NECTAR BABY MANZANA 124G	
NECTAR BABY MANZANA 124G	8	41275		330200
NECTAR BABY MANZANA 124G	7	41276		288932
NECTAR BABY MANZANA 124G	2	41277		82554
NECTAR BABY MANZANA 124G	3	41278		123834
NECTAR BABY MANZANA 124G	5	41279		206395
NECTAR BABY MANZANA 124G	7	41280		288960
NECTAR BABY MANZANA 124G	9	41281		371529
NECTAR BABY MANZANA 124G	5	41282		206410
NECTAR BABY MANZANA 124G	5	41283		206415

Figura 1. Fracción de la base de datos parte de ventas

Fuente: base de datos de la empresa estudiada.

las ecuaciones [1], [2] y [3], donde se abarca el desarrollo del punto de reorden, inventario de seguridad y cantidad económica de producto.

Cantidad económica de producto:

$$CEP = \sqrt{\frac{2DS}{H}} \quad [1]$$

Inventario de seguridad:

$$IS = z\sigma_L \quad [2]$$

Punto de reorden:

$$R = \bar{d}L + z\sigma_L \quad [3]$$

Al aplicar este análisis matemático, se observaron los siguientes resultados (figuras 2 y 3).

Por producto

NECTAR MANZANA X 237 BOT X 6

Histograma Pastel

Dispersión Linealizar

Previsión

Fecha: 13/07/2013

Predecir 6.0 Paquetes

Fuzzy 4.0 Paquetes

Promedio de ventas: 6.0 Paquetes

Intervalos de confianza: 3 - 9 Paquetes

CEP: 32 Paquetes

Inventario de seguridad: 6 Paquetes

Punto de reorden: 12 Paquetes

Figura 2. Métodos de control de inventario para el producto seleccionado

Fuente: elaboración propia.

En la figura 2 puede apreciarse que la política de inventario más adecuada de acuerdo con el *software* es ordenar al proveedor 6 unidades cada vez que el inventario llegue a 12 de estas, a fin de para garantizar el funcionamiento adecuado del modelo planteado.

En la figura 3 puede apreciarse que la política de inventario más adecuada, de acuerdo con el *software* para el producto Néctar Manzana \times 237 Bot. \times 6, es ordenar al proveedor 32 unidades cada vez que el inventario llegue a 12 de estas, a fin de garantizar el funcionamiento adecuado del modelo planteado.

Por producto

NECTAR MANZANA X 237 BOT X 6

Histograma Pastel

Dispersión Linealizar

Previsión

Fecha: 13/07/2013

Predecir 6.0 Paquetes

Fuzzy 4.0 Paquetes

Promedio de ventas: 6.0 Paquetes

Intervalos de confianza: 3 - 9 Paquetes

CEP: 32 Paquetes

Inventario de seguridad: 6 Paquetes

Punto de reorden: 12 Paquetes

Figura 3. Métodos de control de inventario para el producto Néctar Manzana × 237 Bot. × 6

Fuente: elaboración propia.

Modelo basado en la lógica fuzzy como técnica de minería de datos

La figura 4 muestra el diagrama general del proceso que se utilizó para aplicar la lógica *fuzzy* al problema planteado por la empresa, para la administración adecuada del inventario.

Como se muestra en la figura 4, se aplica en un primer momento el proceso de *fuzzificación*, que consiste en convertir una variable real en un grado de pertenencia que cuantifica el grado de posesión hacia su correspondiente variable lingüística. Luego se aplica una *borrosificación* o *fuzzificación* para establecer las entradas del sistema (ventas realizadas e inventario actual), con las cuales hay reglas para la toma de decisiones. Para determinar la función de membresía, se utilizó en la etapa de *fuzzificación* el comportamiento observado con el *software* (figuras 5 y 6).

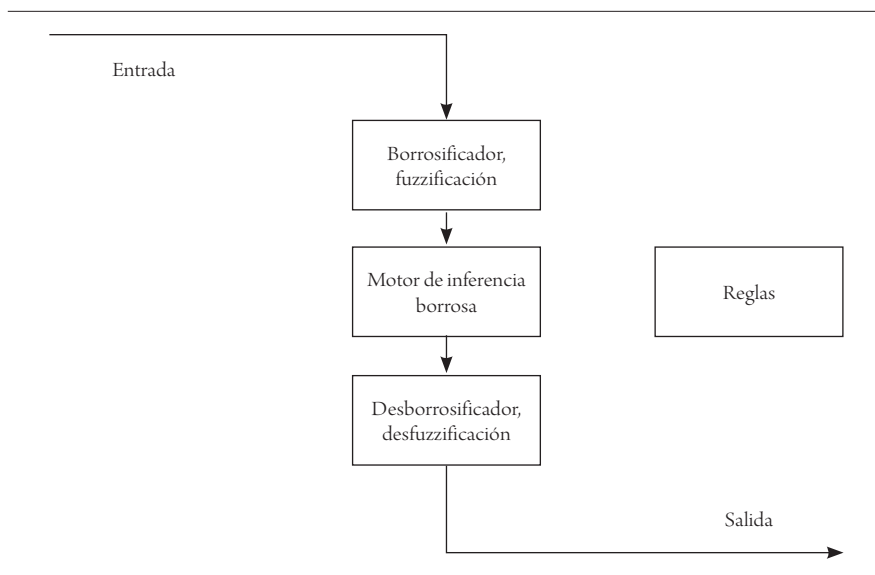
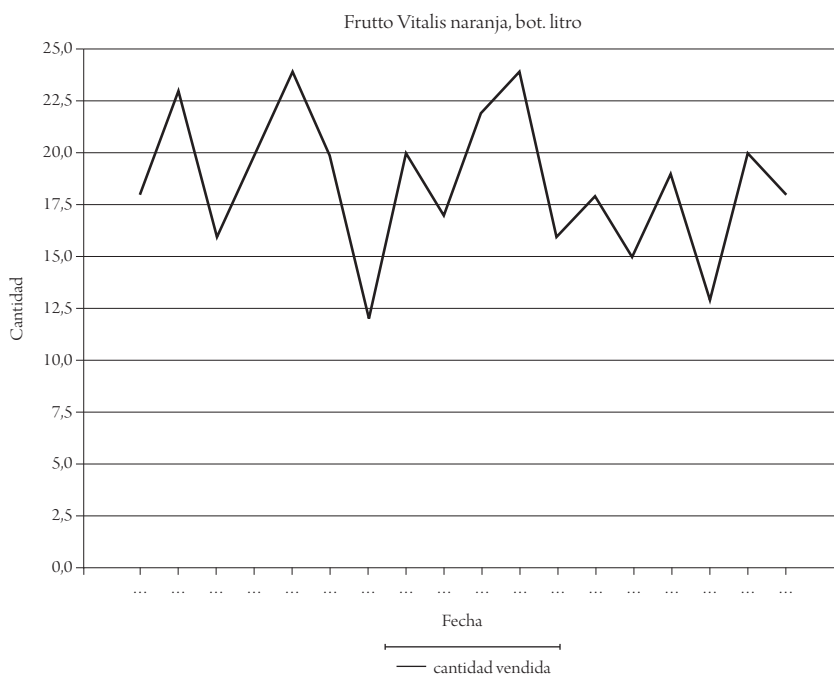


Figura 4. Diagrama de un sistema lógico borroso

Fuente: elaboración propia.



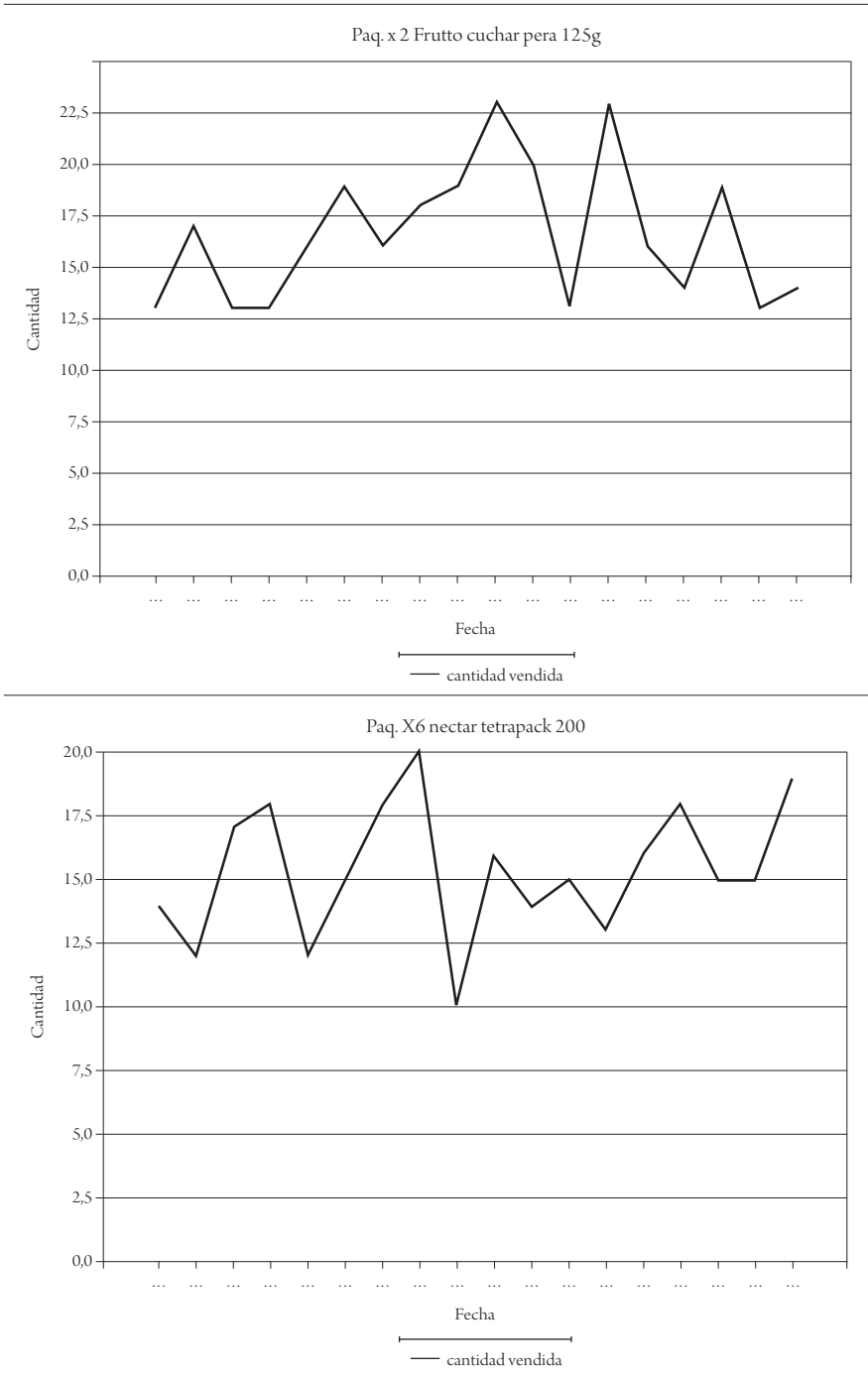


Figura 5. Comportamiento descubierto con el *software*

Fuente: elaboración propia.

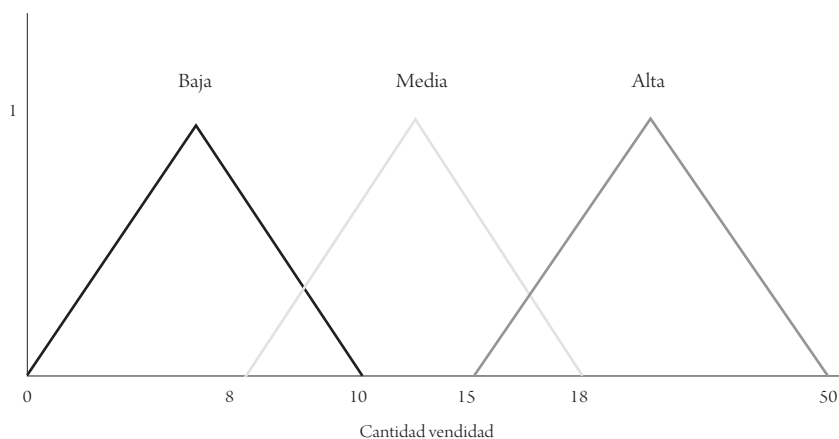


Figura 6. Función de membresía para la variable *cantidad vendida*

Fuente: elaboración propia.

Para este caso, se tuvieron en cuenta principalmente dos variables lingüísticas: *cantidad vendida* y *baja, media y alta*. Cada término es caracterizado por un conjunto difuso dentro de un conjunto universo de los posibles valores. Al asumir que el conjunto universo es *cantidad vendida* y está en el rango (0,50 unidades), entonces los términos lingüísticos son: bajo, 0-10 unidades; medio, 8-18 unidades; alto, 15-50-unidades. De igual forma se realiza con la variable *stock* (figura 7).

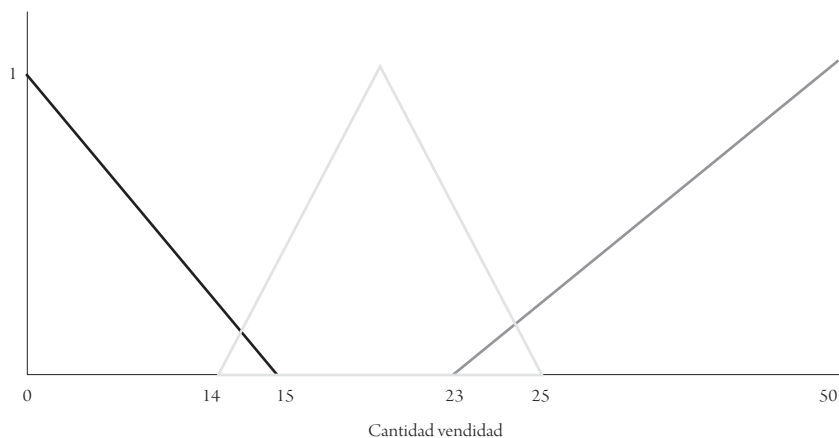


Figura 7. Función de membresía para la variable *stock*

Fuente: elaboración propia.

Una vez es establecida la función de pertenencia para cada una de las variables, como se hizo en el paso anterior, se generaron las reglas específicas (figura 8).

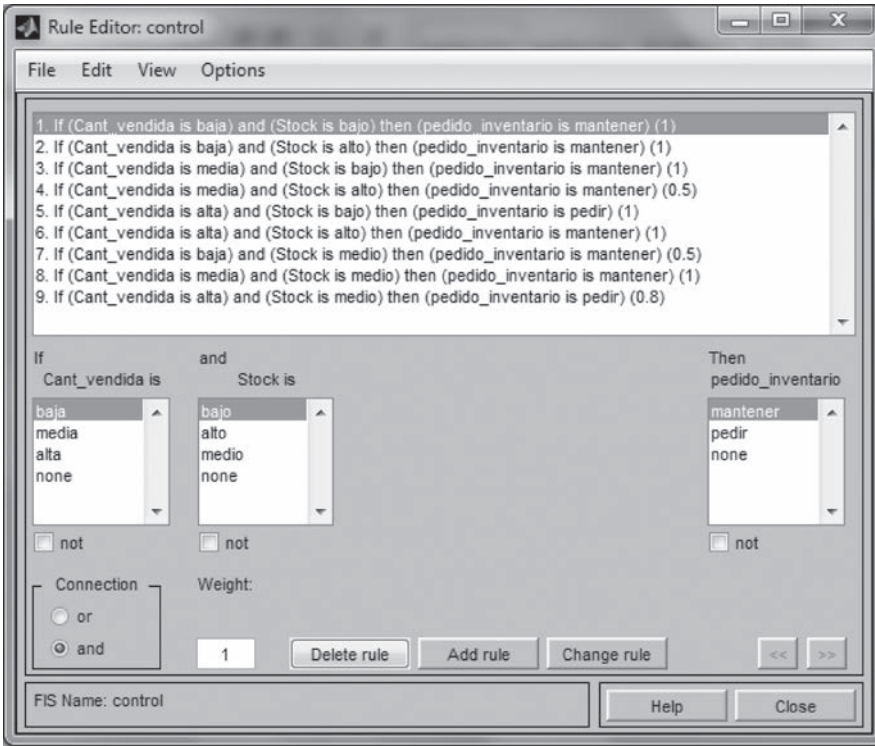


Figura 8. Motor de reglas del sistema

Fuente: elaboración propia.

Cuando se cumple la regla número 1, donde la cantidad vendida es baja y el *stock* también lo es, la respuesta de salida es mantener y no hacer pedidos, a fin de conservar el inventario (figura 9). La figura 10 muestra la relación en 3D de las variables y la respuesta del sistema. Se pueden ver de manera integrada las reglas del motor de inferencia, donde si el *stock* está por debajo de 20 y la cantidad vendida está por encima de 20-23 unidades, entonces el pedido para abastecer el inventario tiene una elevación mayor, o viceversa.

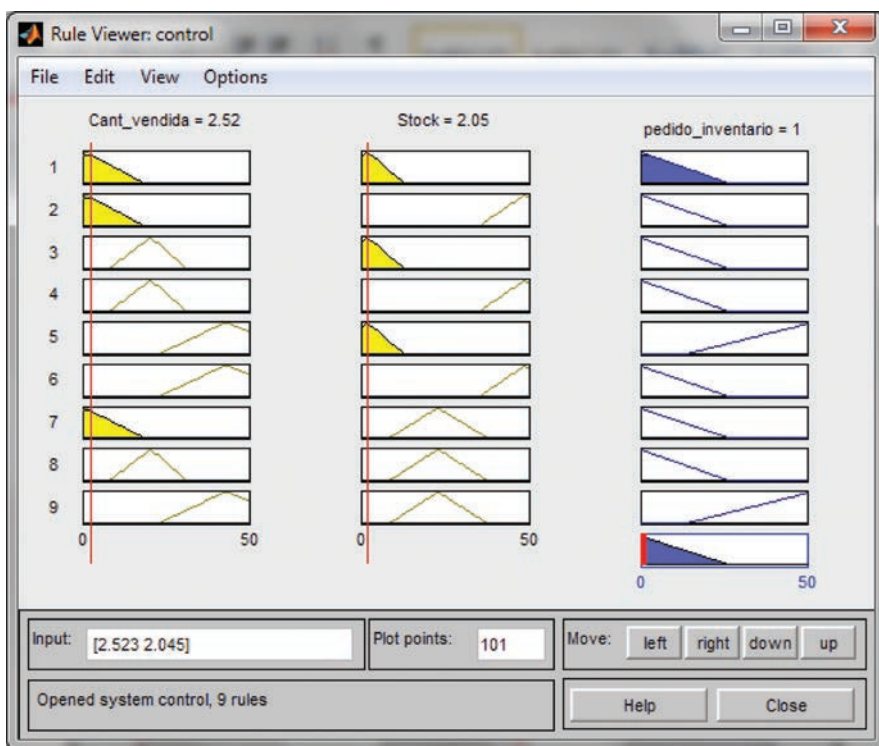


Figura 9. Regla 1: defuzzificación

Fuente: elaboración propia.

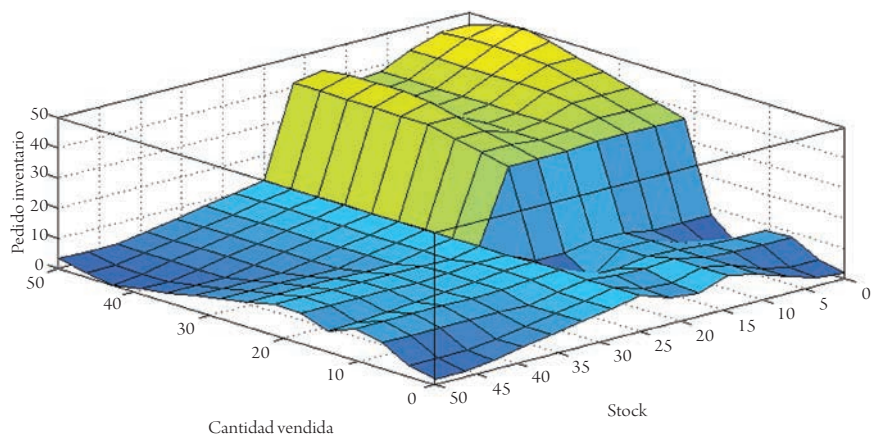


Figura 10. Representación de las reglas en 3D

Fuente: elaboración propia.

Metodología de desarrollo de software

La metodología para el desarrollo de *software*, también denominada *ciclo de vida del desarrollo de sistemas de información*, consta de los procedimientos que se observan en la figura 11.

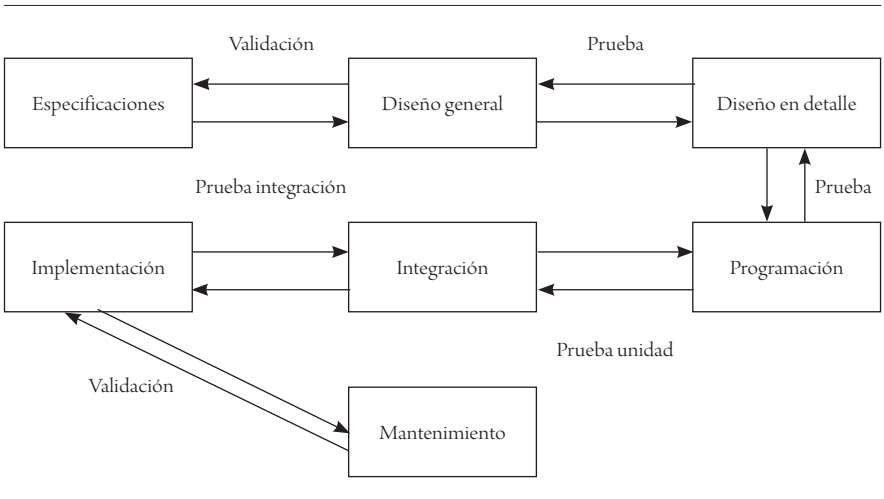


Figura 11. Etapas del ciclo de vida de *software*

Fuente: elaboración propia.

En la definición de los requerimientos de los sistemas es importante establecer los casos de uso, como se observa en las figuras 12 y 13.

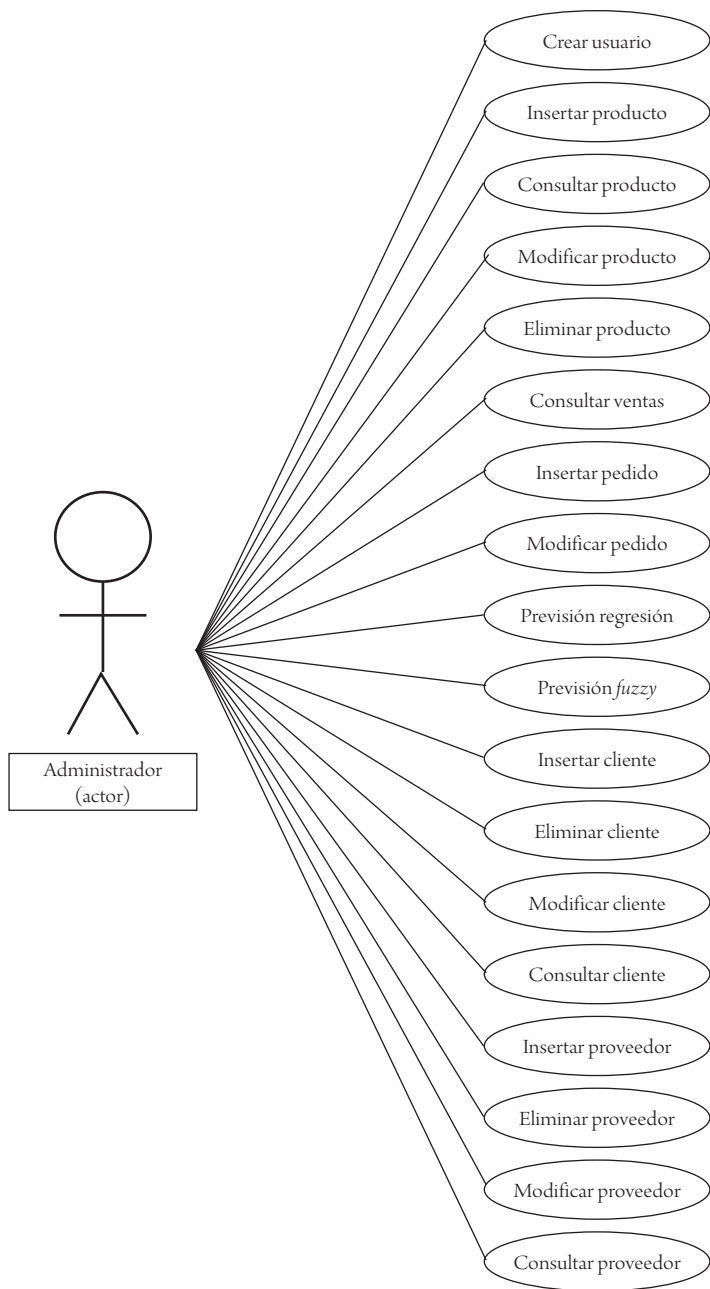


Figura 12. Casos de uso para el administrador

Fuente: elaboración propia.

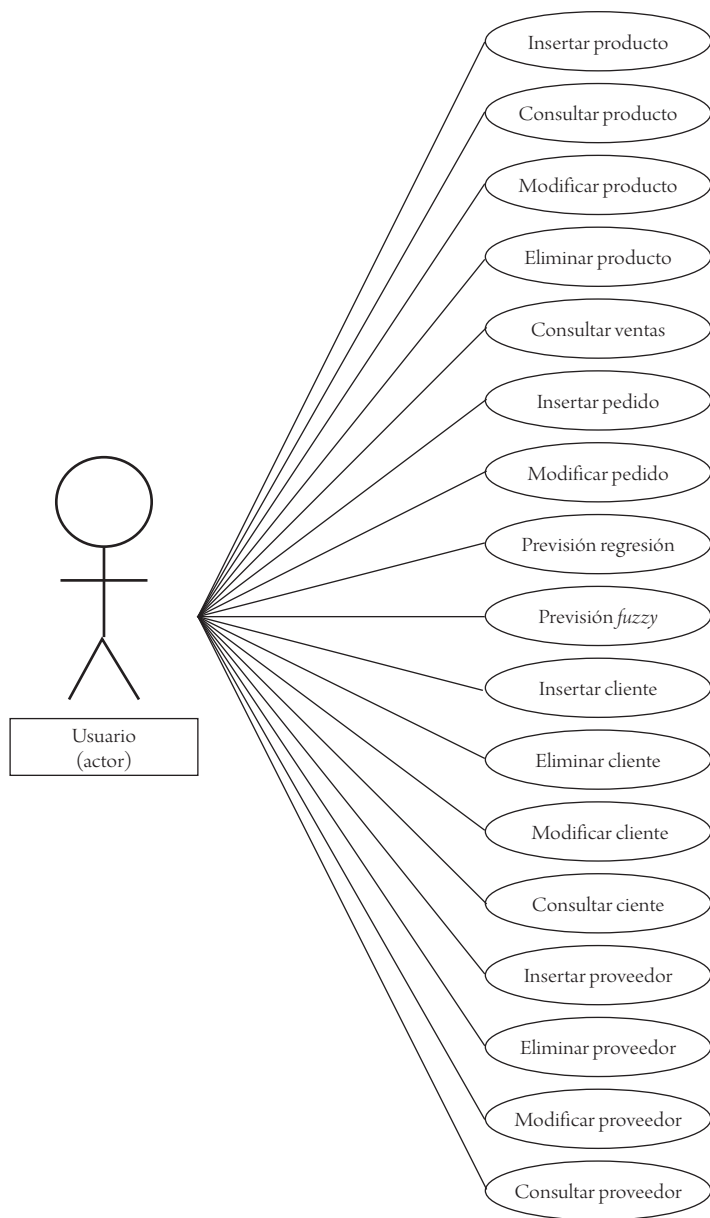


Figura 13. Casos de uso para el usuario

Fuente: elaboración propia.

Una vez definidos los requerimientos, se diseñó la interfaz gráfica, cuyos componentes se muestran en la figura 14.

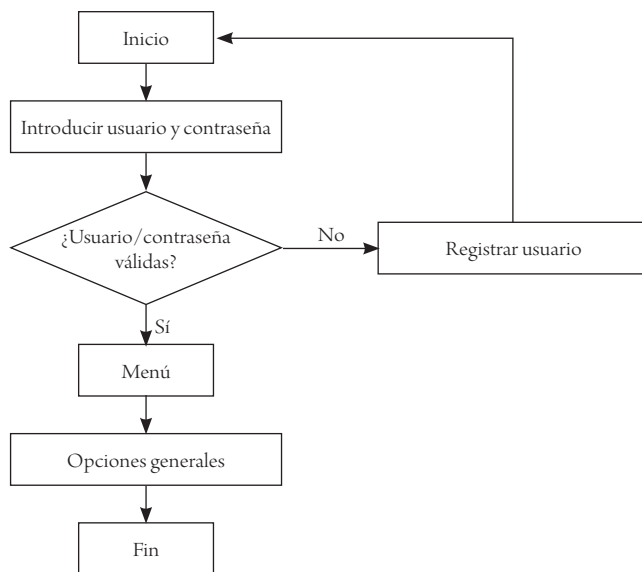


Figura 14. Diagrama de acceso al sistema

Fuente: elaboración propia.

La figura 15 corresponde al esquema general para desplazarse a través de las diferentes opciones que tiene el sistema, una vez se haya accedido a él.

Resultados

Con la librería JFreeChart y a partir de los datos del sistema pueden obtenerse diversos gráficos para el análisis de las ventas de los productos de la empresa. La figura 16 muestra las ventas totales correspondientes al periodo 1-20 de enero de 2013, donde se compara cantidad contra producto. De esta manera, se obtiene una descripción más agradable de las ventas de un producto que la simple consulta de la tabla de la base de datos.

Posteriormente, en las figuras 17 y 18 pueden apreciarse claramente las ventas de un producto determinado mediante un gráfico de dispersión. En este caso, el producto seleccionado es Néctar Baby Manzana 124G. Es importante destacar la comparación entre la cantidad vendida del producto seleccionado con el rango de fechas que se encuentran en la base de datos. Luego se hace una comparación con Excel para determinar que el sistema sí grafica de forma correcta la información.

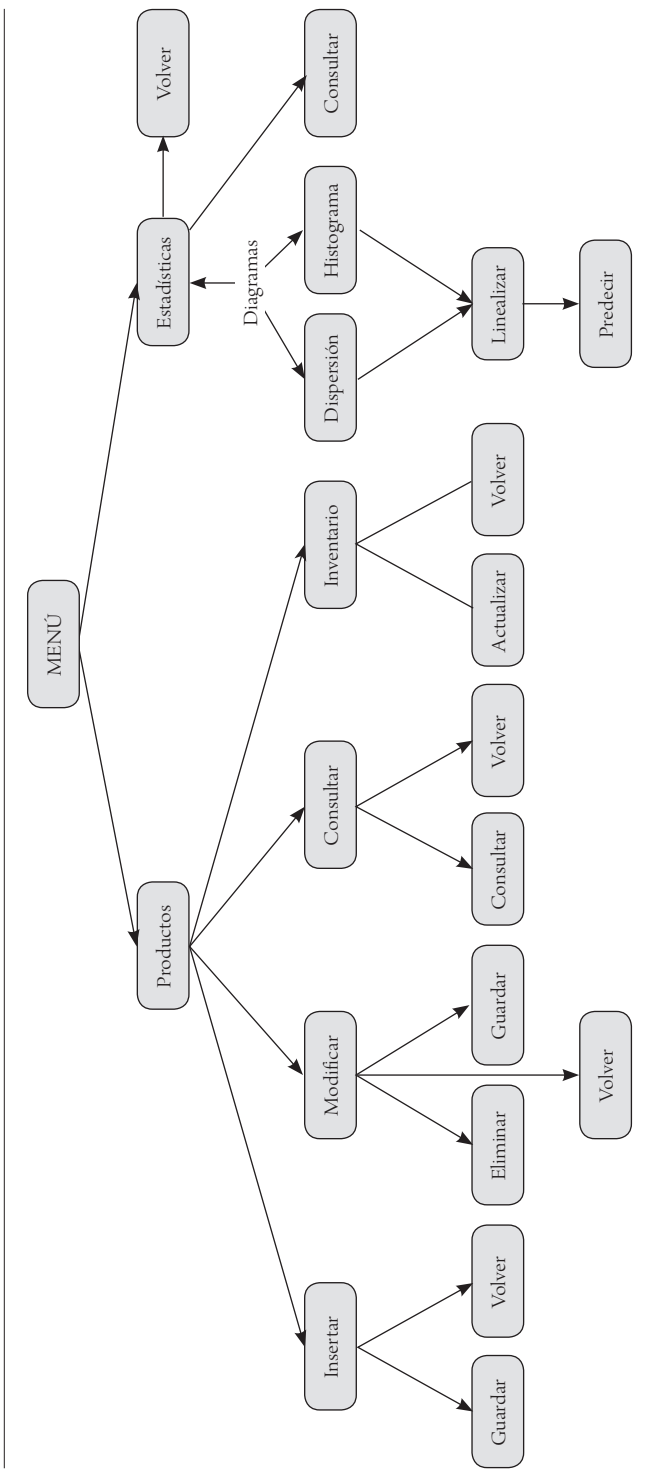


Figura 15. Mapa de navegación

Fuente: elaboración propia.

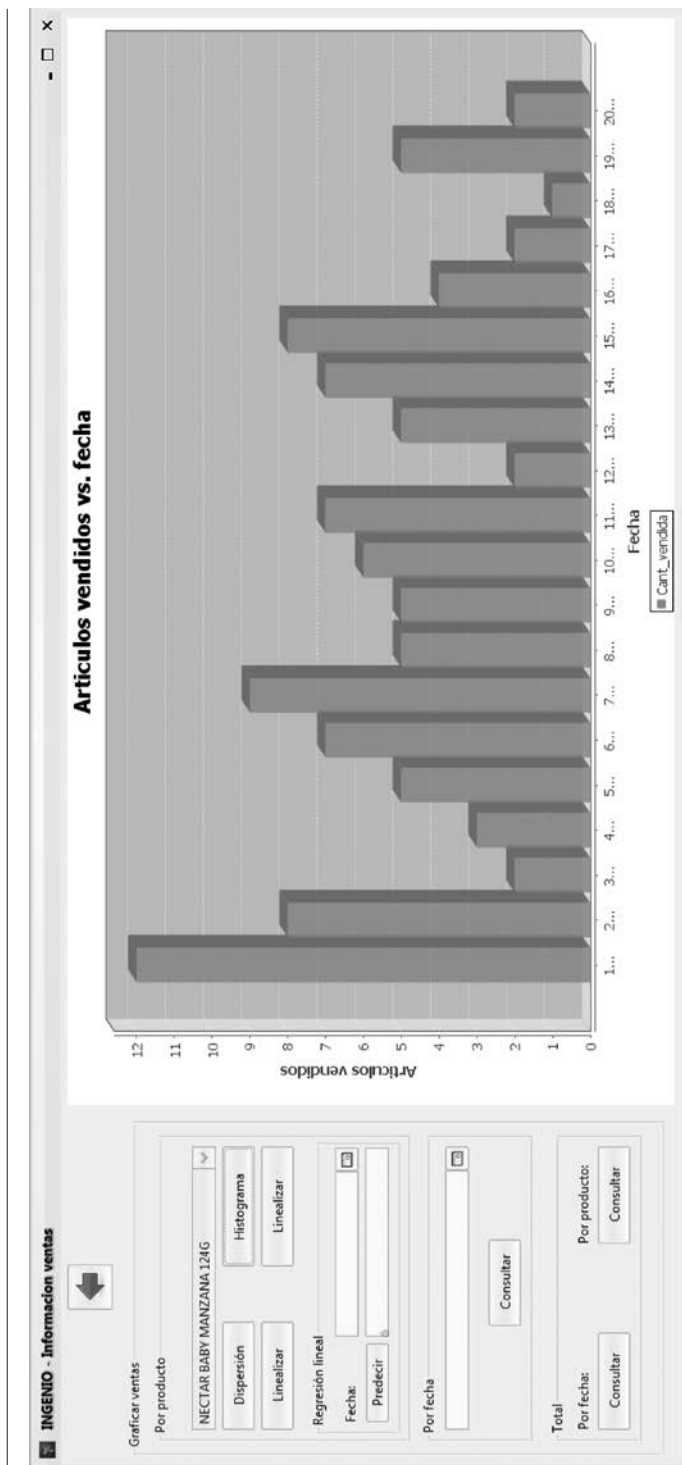


Figura 16. Ventas correspondientes del 1 al 20 de enero de 2013

Fuente: elaboración propia.

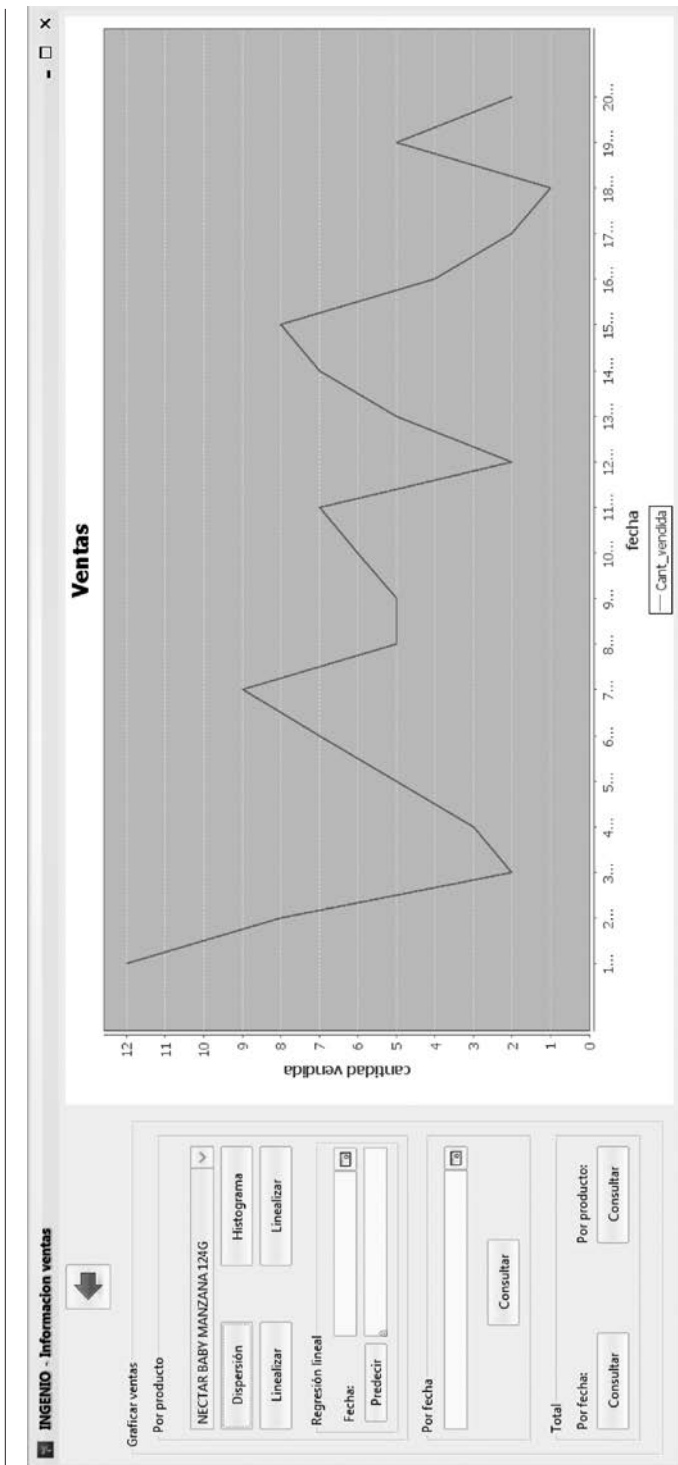


Figura 17. Dispersión para las ventas de Néctar Baby Manzana 124G obtenido con el sistema

Fuente: elaboración propia.

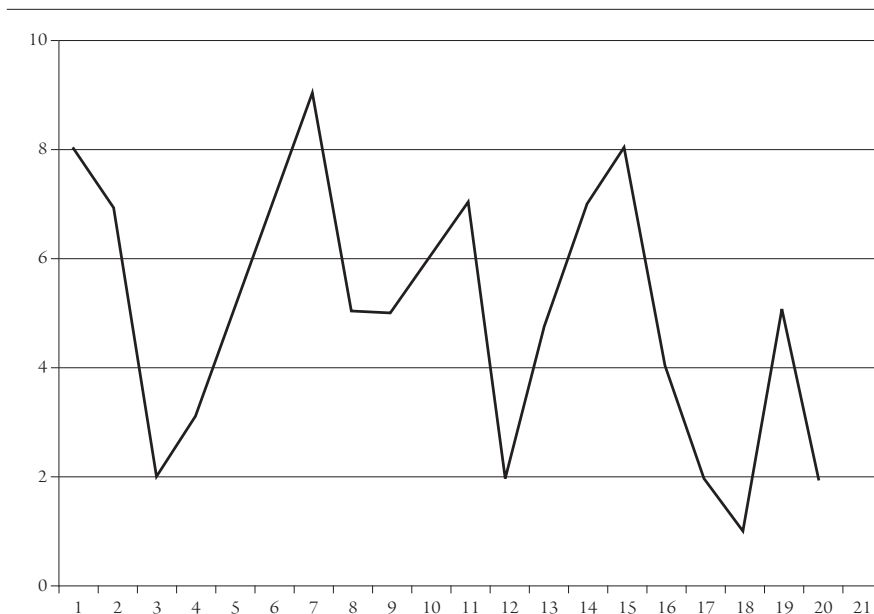


Figura 18. Dispersión para las ventas de Néctar Baby Manzana 124G obtenido con Excel

Fuente: elaboración propia.

En las figuras 19 y 20 puede observarse de manera distinta el comportamiento mediante un gráfico de barras. En este caso también se hace la comparación entre el sistema y Excel.

Para realizar la predicción de ventas a través del método matemático de regresión lineal, es necesario implementar un algoritmo que le permita al usuario interactuar de manera adecuada y sencilla entre la predicción deseada y la aplicación. En primer lugar, se realizan todas las operaciones de prueba en Excel, de donde se parte para la obtención del modelo. Para obtener los valores que forman parte de la ecuación general de la recta es necesario determinar los coeficientes b y C que se muestran a continuación.

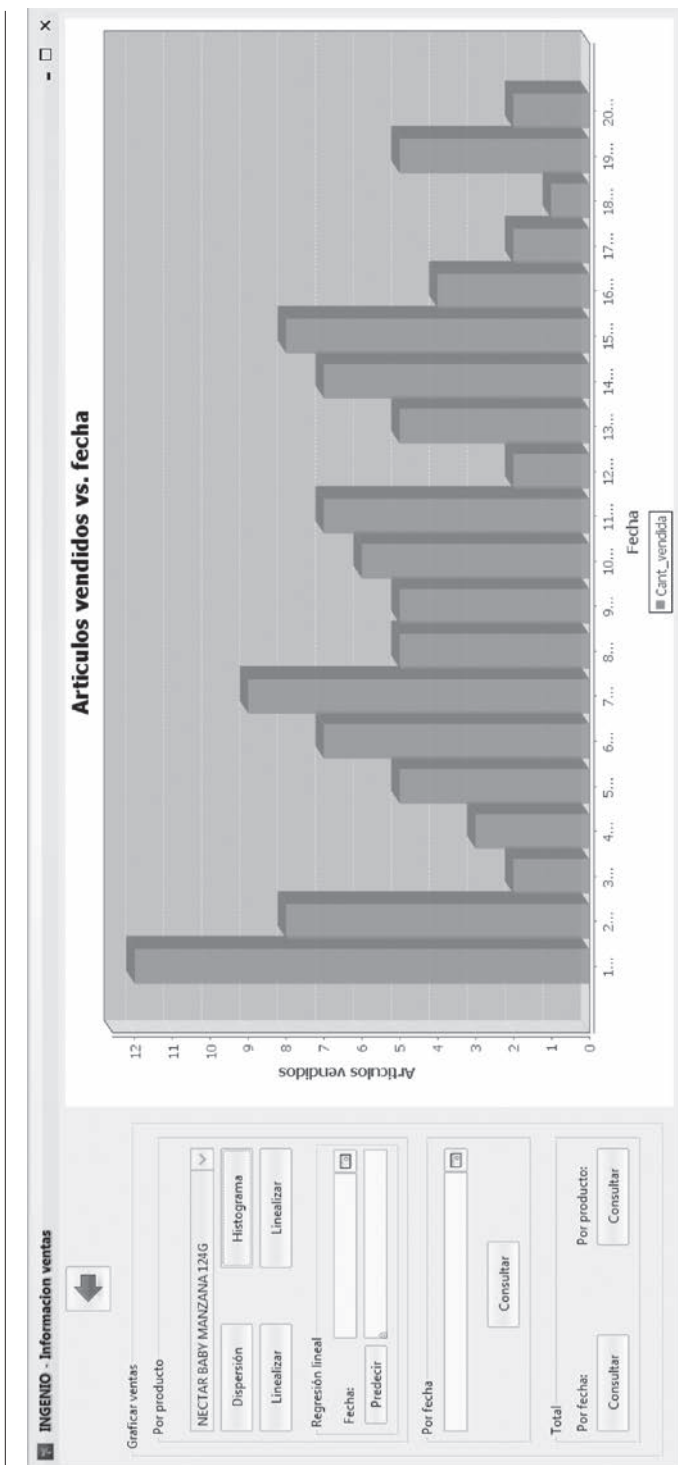


Figura 19. Ventas de Néctar Baby Manzana 124G en el sistema

Fuente: elaboración propia.

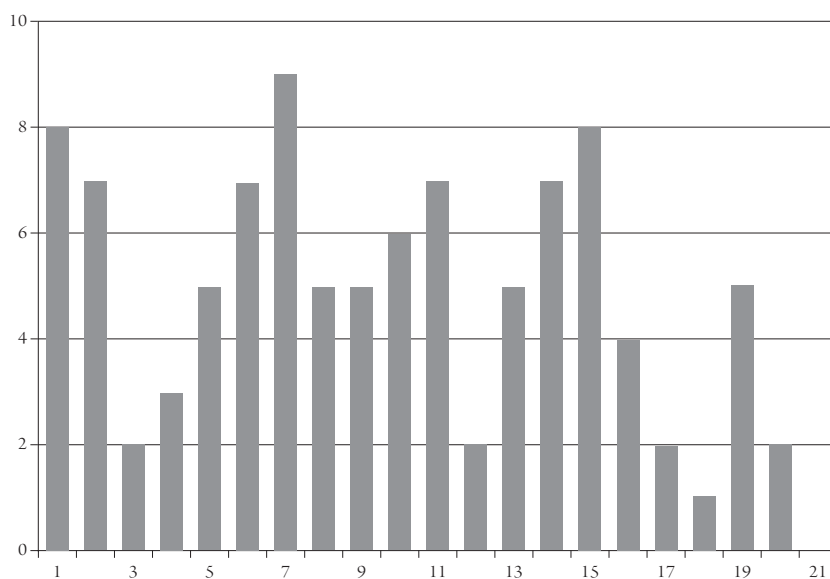


Figura 20. Ventas de Néctar Baby Manzana 124G en Excel

Fuente: elaboración propia.

Ecuación general de la recta:

$$\hat{Y} = bX + C$$

Ecuación de coeficiente b :

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

Ecuación de coeficiente C :

$$C = \bar{Y} - b\bar{X}$$

En las ecuaciones anteriores se denotan claramente las diversas variables, donde X es la variable independiente (para este caso, la fecha de venta del producto); Y es la variable dependiente (la cantidad de productos vendidos en el día); \bar{X} es la

media aritmética de la variable independiente, y \bar{Y} es la media aritmética de la variable dependiente.

Para la implementación del modelo de regresión lineal, en primer lugar fue necesario convertir la fecha de venta del producto en un número, para lo cual se tomó el algoritmo utilizado por Microsoft Excel, que establece que el día 0 es el 1 de enero de 1900. Esto permite definir y manejar el algoritmo de manera más sencilla y veloz. Igualmente fue necesario almacenar los datos de los coeficientes b y C en la base de datos, a fin de favorecer la implementación del algoritmo y guardar los puntos de cada dato en la tabla *am_sales_lineal* para el posterior análisis del usuario. Ello puede realizarse fácilmente accediendo al panel de regresión lineal de la ventana de ventas.

Para la validación de los datos obtenidos a través de la implementación del algoritmo expuesto anteriormente fue necesario realizar la validación de los datos en Excel, donde se implementó el algoritmo, como se observa en las siguientes tablas para las ventas del producto Frutto Clasico Pera (caja litro), entre el 1 y el 20 de enero de 2013. Como ya existe una serie de datos históricos en el sistema de información, se decidió hacer algunas pruebas con diferentes fechas hacia futuro, cuyos resultados se muestran en las tablas 2, 3 y 4.

Tabla 2. Declaración de variables y factorización de los elementos de la ecuación para obtener los coeficientes b y C de la ecuación de la recta

X	Y	X*Y	X^2
FECHA	PRODUCTO		
41275,00	5	206375	1703625625
41276,00	2	82552	1703708176
41277,00	6	247662	1703790729
41278,00	3	123834	1703873284
41279,00	5	206395	1703955841
41280,00	2	82560	1704038400
41281,00	3	123843	1704120961
41282,00	5	206410	1704203524
41283,00	7	288981	1704286089
41284,00	9	371556	1704368656

Continúa

X	Y		
FECHA	PRODUCTO	X*Y	X^2
41285,00	11	454135	1704451225
41286,00	10	412860	1704533796
41287,00	7	289009	1704616369
41288,00	9	371592	1704698944
41289,00	5	206445	1704781521
41290,00	3	123870	1704864100
41291,00	8	330328	1704946681
41292,00	7	289044	1705029264
41293,00	7	289051	1705111849
41294,00	6	247764	1705194436

Fuente: elaboración propia.

Tabla 3. Coeficientes *b* y *c* obtenidos a través de Excel

X	Y	SUM (X * Y) – N * X * Y	SUM (X ^ 2) – N(X ^ 2)	B	C
41284,5	6	4458852	30679379590	0,000145337	-0,00016943

Fuente: elaboración propia.

Tabla 4. Valores obtenidos para la linealización del producto escogido

LINEALIZACIÓN
5,9986193
5,99876463
5,99890997
5,99905531
5,99920065
5,99934598
5,99949132
5,99963666
5,99978199
5,99992733
6,00007267
6,00021801

LINEALIZACIÓN
6,00036334
6,00050868
6,00065402
6,00079935
6,00094469
6,00109003
6,00123537
6,0013807

Fuente: elaboración propia.

Los valores anteriores coinciden adecuadamente con los obtenidos en la base de datos, como se puede corroborar con la tabla 5.

Tabla 5. Tabla *am_sales_lineal* con los coeficientes *b* y *C* para el producto Frutto Clásico Pera (caja litro)

X_FECHA	X_FECHA_NUMERO	Y_LINEAL	B	C
1-1-2013	41275	5.99861929756837	0.000145337098065896	-0.000169425101470111
2-1-2013	41276	5.99876463466644	0.000145337098065896	-0.000169425101470111
3-1-2013	41277	5.99890997176451	0.000145337098065896	-0.000169425101470111
4-1-2013	41278	5.99905530886257	0.000145337098065896	-0.000169425101470111
5-1-2013	41279	5.99920064596064	0.000145337098065896	-0.000169425101470111
6-1-2013	41280	5.9993459830587	0.000145337098065896	-0.000169425101470111
7-1-2013	41281	5.99949132015677	0.000145337098065896	-0.000169425101470111
8-1-2013	41282	5.99963665725484	0.000145337098065896	-0.000169425101470111
9-1-2013	41283	5.9997819943529	0.000145337098065896	-0.000169425101470111
10-1-2013	41284	5.99992733145097	0.000145337098065896	-0.000169425101470111
11-1-2013	41285	6.00007266854903	0.000145337098065896	-0.000169425101470111
12-1-2013	41286	6.0002180056471	0.000145337098065896	-0.000169425101470111
13-1-2013	41287	6.00036334274516	0.000145337098065896	-0.000169425101470111
14-1-2013	41288	6.00050867984323	0.000145337098065896	-0.000169425101470111
15-1-2013	41289	6.0006540169413	0.000145337098065896	-0.000169425101470111
16-1-2013	41290	6.99979935403936	0.000145337098065896	-0.000169425101470111

Continúa

X_FECHA	X_FECHA_ NUMERO	Y_LINEAL	B	C
17-1-2013	41291	6.00094469113743	0.000145337098065896	-0.000169425101470111
18-1-2013	41292	6.00109002823549	0.000145339098065896	-0.000169425101470111
19-1-2013	41293	6.00123536533356	0.000145337098065896	-0.000169425101470111
20-1-2013	41294	6.00138070243163	0.000145337098065896	-0.000169425101470111

Fuente: elaboración propia.

Para realizar la validación de pruebas, en este caso se eligió como ejemplo el día 29 de marzo de 2013, el cual obtuvo un valor predictivo de 6,01126 mediante la implementación del algoritmo en Excel. Este dato coincide correctamente con el valor que muestra el *software* (figura 21).

Conclusiones

La aplicabilidad de la minería de datos en sistemas de gestión depende en gran medida de la herramienta y del tratamiento de los datos que se requieran analizar. Para este caso, es necesario aclarar que los valores obtenidos no brindan una descripción adecuada de los datos proporcionados por el cliente; sin embargo, la minería de datos puede contribuir significativamente en las aplicaciones de administración empresarial basada en la relación con el cliente. Debe tenerse en cuenta, en este sentido, que a pesar de que la herramienta elegida en esta ocasión no brindó los resultados esperados, sí ayudó a confirmar que la dispersión de los datos puede llevar a utilizar métodos más comunes, como la linealización de las gráficas de las ventas obtenidas.

Los requerimientos del sistema de información se evidencian claramente en cada uno de los formularios desarrollados, en los que se tienen en cuenta las necesidades principales del cliente, como la consulta de la base de datos de los productos, la modificación de estos, el manejo robusto del inventario de la compañía y el análisis descriptivo a través de diferentes gráficas de las ventas de los productos utilizando diferentes parámetros (por ejemplo, un producto determinado o un rango de fechas escogido por el usuario del sistema). De igual manera, existen dentro del mismo sistema requerimientos básicos de un sistema de información de transacciones como el registro de usuarios y el ingreso de estos utilizando un nombre de usuario y una contraseña.

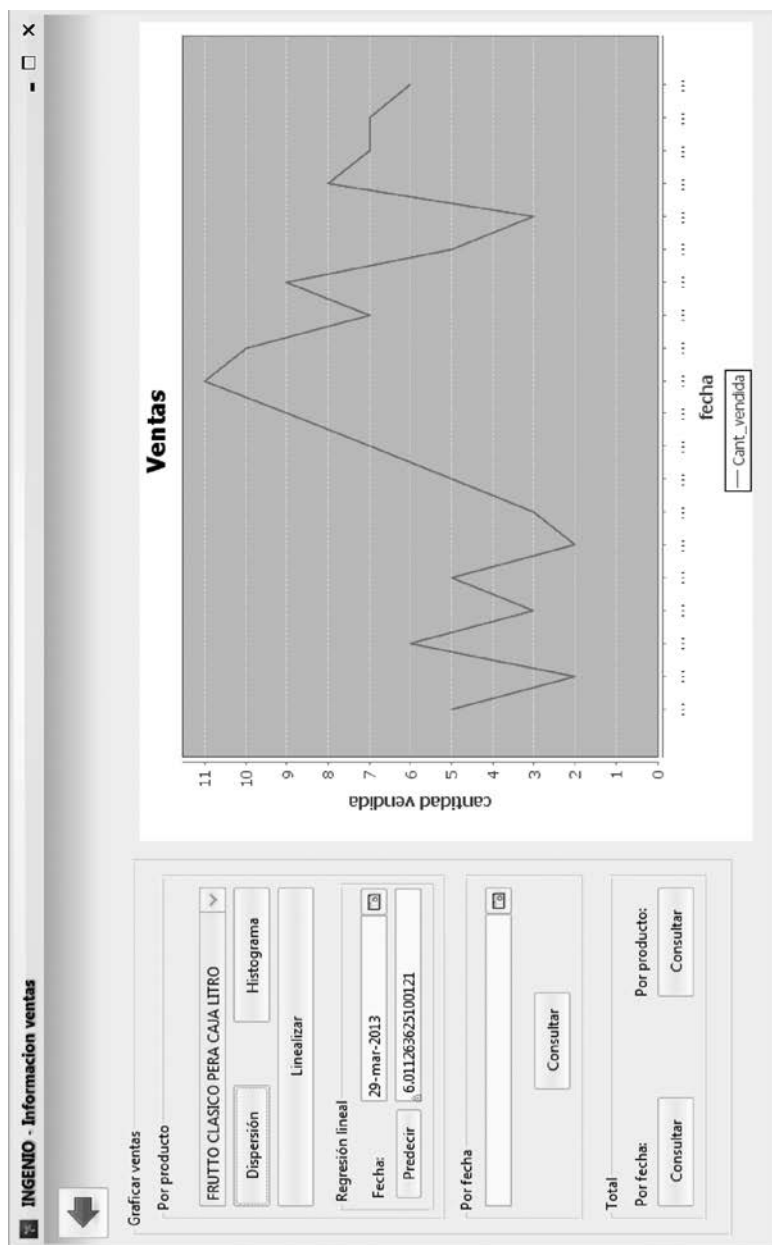


Figura 21. Predicción de ventas para una fecha determinada y el artículo seleccionado

Fuente: elaboración propia.

El desarrollo del diseño del sistema de información no solo implicó un análisis profundo de los requerimientos del sistema, sino que además incluyó el estudio de una metodología de minería de datos en conjunto con herramientas adicionales para la administración de la base de datos, así como el desarrollo y la descripción gráfica de la información contenida dentro de las bases de datos de la empresa.

Finalmente, a través de la lógica *fuzzy* se observó que las entradas y salidas pueden relacionarse sin necesidad de entender todas las variables involucradas en el proceso. Ello permite que el sistema pueda ser más confiable y estable que uno con un sistema de control convencional. Es posible obtener prototipos rápidamente, ya que no requiere conocer todas las variables acerca del sistema antes de empezar a trabajar, por cuanto su desarrollo es más económico que el de sistemas convencionales, porque son más fáciles de designar.

Referencias

- Alcántara Mori, Á. A. (2012). *Formulación de minería de datos para la empresa distribuidora de productos Espinoza Aguilar S.A.* (tesis de grado). Universidad Tecnológica del Perú, Lima, Perú.
- Elizondo Gracia, I. (2002). *Desarrollo metodológico y técnico de un pronóstico de ventas para un producto con patrones cíclicos* (tesis de grado). Universidad Autónoma de Nuevo León, Monterrey, México.
- Fortuño, F. J. (2009). *Desarrollo y evolución de un sistema de gestión de inventarios, ventas y facturación para una empresa automotriz* (tesis de grado). Universidad de Chile, Santiago de Chile, Chile.
- García Pérez, R. D. (2009). *Desarrollo de un software para el control de inventario de productos terminados para los departamentos de atención al cliente, la línea de producción "sector beta" y despacho en una empresa alimentos* (tesis de grado). Universidad de Oriente, Cumaná, Venezuela.
- Luna Beltrán, G. A. y Quiroz Morales, D. F. (2009). *Diseño e implementación de un sistema ERP para optimizar sistemas de gerencia de restaurantes o bares protegido por una llave electrónica* (tesis de grado). Universidad de La Salle, Bogotá, Colombia.
- Moreno Malaver, L. A. (2005). *Diseño e implementación de un sistema de información modular para la compra venta e inventario de una fábrica de calzado* (tesis de grado). Universidad de La Salle, Bogotá, Colombia.

- Perversi, I. (2007). 'Aplicación de minería de datos para la exploración y detección de patrones delictivos en 2007 en Argentina. Recuperado de <http://iidia.com.ar/rgm/tesis/PERVERSI-tesisdegradoeningenieria.pdf>
- Ropero Rodríguez, J. (2009). *Método general de extracción de información basado en el uso de lógica borrosa. Aplicación en portales web* (tesis de grado). Sevilla: Universidad de Sevilla.
- Vallejos, S. J. (2006). *Minería de datos* (tesis de grado). Corrientes, Argentina: Universidad Nacional del Nordeste, Corrientes, Argentina.

