

APLICACIÓN DE BUSINESS INTELLIGENCE BASADO EN UN MODELO DE PRONÓSTICO DE SERIES DE TIEMPO PARA MEJORAR EL VOLUMEN DE VENTAS DE LA EMPRESA FSRL DE LIMA

Autor: Juan Castro¹

Asesor: Mg. Sergio Valladares Castillo

e-mail: 1juanp@upeu.edu.pe

Resumen

En la empresa de panificación FSRL las decisiones no son las más adecuadas debido a que la información de valor no está disponible en el tiempo requerido ni con la calidad suficiente. El presente artículo tiene como objetivo desarrollar un sistema de inteligencia de negocios basado en un modelo de pronóstico de ventas en FSRL. La solución consiste en desarrollar un Data mart de ventas bajo la metodología kimball. El sistema ofrece reportes inteligentes y un pronóstico de ventas basado en modelos estocásticos con un mínimo porcentaje de error. La empresa evaluará en un periodo determinado el sistema. Se proyecta obtener pronósticos lo más reales posibles para la buena toma de decisiones de FSRL.

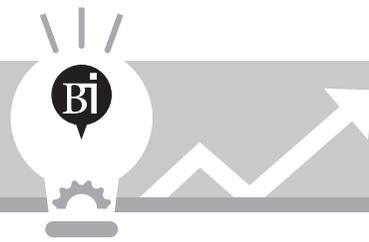
Índice de términos—BI: inteligencia de negocios, DSS: sistema de apoyo a las decisiones, EIS: sistema de información para ejecutivos.

I. INTRODUCCIÓN

La empresa de panificación FSRL cuenta con una base de datos operacional, “Figueri”, útil para el trabajo operativo de ventas, pero no para el nivel estratégico, que necesita información procesada, útil y disponible para tomar decisiones de calidad. La gerencia necesita información de ventas y pronóstico de ventas por los diferentes tipos y líneas de productos, conociéndose así la tendencia de las ventas de cada producto y así tomar decisiones de acuerdo al ciclo de vida del producto. El encargado del área de sistemas utiliza excesivo tiempo para recopilar información de la base de datos y transformarlo en un informe de ventas adecuado a la gerencia y al área de finanzas. El tiempo y el costo para realizar este trabajo son considerables en el caso de que la gerencia pida un informe detallado de varios meses. Este problema con el tiempo ha generado demora en la toma de decisiones debido a que se utiliza herra-

mientas manuales y como consecuencia no hay una actuación rápida cuando un producto cae en el mercado.

En la actualidad la Inteligencia de negocios en el mercado ha ganado mucha acogida debido a la necesidad de los empresarios en el nivel estratégico, y la importancia que tiene para resolver problemas estratégicos. La toma de decisiones es un proceso complejo que requiere el apoyo de todas las herramientas disponibles[1]. La organización empresarial tiene que analizar el mercado de manera que sea capaz de permanecer estable, y hacer frente a los cambios del mercado y finalmente ser capaz de manejar la gestión del mercado. Para ello, la organización debe actualizar sus procesos de negocio mediante la utilización de tecnologías modernas, como la inteligencia de negocios (BI). Es decir que la implementación de la inteligencia de negocios brinda beneficios como el aumento de la seguridad en la toma de decisiones estraté-



gicas, modificar las múltiples comunicaciones en la organización y aumentar el conocimiento del mercado [2]. A lo largo de la última década la inteligencia de negocios ha evolucionado, y cada vez más se requiere de datos en tiempo real. Con BI se puede generar reportes globales o por secciones, crear escenarios con respecto a una decisión y generar pronósticos de ventas y compras [3]. Lo que la empresa FRLS necesita en el nivel estratégico es la construcción de un Data Mart de ventas para el área administrativa.

Los sistemas de apoyo a las decisiones (DSS) son algunas de las nuevas aplicaciones computarizadas que actúan como un sistema de apoyo para ayudar a los administradores a monitorear el desempeño operativo o ganar “inteligencia” de los datos históricos[4]. Un DSS eficaz puede compilar la información más importante de los documentos, modelos de negocio y datos brutos e incluso ayudar a resolver problemas y tomar decisiones útiles mediante la utilización de diversas herramientas OLAP. Consiste en la integra-

ayudan a analizar, comparar y resaltar las tendencias de las variables importantes a los ejecutivos de alto nivel para que puedan monitorear el desempeño e identificar las oportunidades y problemas [6] y [7].

Un Data mart es un subconjunto de un Data warehouse, donde una parte reducida que almacena un subconjunto específico de los datos de una compañía es ubicada en una base de datos separada para uso personal específico[8], es decir utilizado por un departamento individual. Un Data warehouse contiene datos históricos integrados. Desde el Data warehouse fluyen datos a los distintos departamentos para su uso en su sistema personalizado de soporte de decisiones (DSS). Según Ha y Park [9] estas bases de datos o Data mart son conocidos como Data warehouse departamental. Los datos que residen en el Data warehouse están en un nivel muy granular y los datos en el Data mart están en un nivel refinado. Hay razones organizativas, tecnológicas y económicas de por qué el mercado de datos se convierte en atractivo así como el Data warehouse crece en volumen.

El método de series de tiempo utiliza datos históricos para estimar resultados futuros. En este tipo se encuentran el promedio móvil, la extrapolación, la predicción lineal, la estimación de tendencias y la curva de crecimiento. En este método se deben tomar en cuenta los siguientes factores: factores de ajustes de ventas; son aquellos que influyen en el aumento o disminución de las ventas. Periodicidad es el número de veces en que el producto se vende en un periodo de tiempo. Otro método son los causales o econométricos, se basan en analizar las variables que causan el evento y poder predecir su comportamiento en el futuro. Entre los métodos que pertenecen a este tipo están el análisis de regresión lineal o no lineal y la econometría.

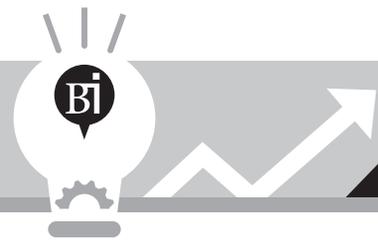
El uso de los métodos de pronósticos es variable, mayormente se combinan varios métodos cuantitativos y cualitativos. Los pronósticos reducen el riesgo de la decisión, usando todas las herramientas disponibles para anticipar el resultado más probable. Para los ejecutivos de las empresas es indispensable para una mejor dirección.

Existen factores que se denominan elementos de cambio, estos son los que afectan a la demanda

ción de datos provenientes de múltiples fuentes y está diseñado para dar solución a los problemas y toma de decisiones mediante la compilación de información de fuentes dispares[5].

Un sistema de información ejecutiva (EIS) es un tipo de sistema de gestión de información destinada a facilitar y apoyar a la toma de decisiones. Es comúnmente considerado como una forma especializada del sistema de apoyo a las decisiones (DSS). El énfasis está en las pantallas gráficas y de uso fácil del interfaces de usuario. Ofrecen informes, alertas y semáforos. En general, los EIS





de las ventas. Si no fuera por estos elementos, los volúmenes de ventas serían una línea horizontal, y se podría saber fácilmente las ventas futuras, básicamente son cinco: (1) la tendencia a largo plazo, (2) el ciclo, (3) fluctuaciones estacionales, (4) variaciones del calendario, (5) acontecimientos irregulares.

Según Nojek, S. [1] un error significativo en el pronóstico de ventas podría dejar a una empresa sin la materia prima o insumos necesarios para su producción, o podría generarle un inventario demasiado grande.

La identificación y estimación de modelos de series temporales han sido desarrolladas para procesos estacionarios, que pueden clasificarse en autorregresivos (AR), de medias móviles (MA) o procesos mixtos (ARMA).

II. METODOLOGÍA

Para el desarrollo del Data mart se adoptó el ciclo de vida dimensional del negocio según [13], que consta de 10 etapas.

A. Planificación

En esta etapa se concentra en identificar la definición y el alcance del proyecto del Data mart basándose en los requerimientos del negocio, y finalmente en el desarrollo del plan del proyecto, seguimiento y la monitorización.

B. Definición de requerimientos

Se realiza la interpretación correcta de los diferentes requerimientos, entendiendo los factores claves que guían el negocio, ya que los requerimientos deben determinar el alcance del Data mart.

C. Diseño de arquitectura técnica

Se debe tener en cuenta tres factores: los requerimientos del negocio, los actuales entornos técnicos y las directrices técnicas y estratégicas futuras planificadas por la compañía para poder establecer el diseño de la arquitectura técnica del entorno de Data mart.

D. Modelado dimensional

Se empieza por definir una matriz donde se determinará la dimensión de cada indicador y lue-

go se especifican los diferentes grados de detalle dentro de cada concepto del negocio.

E. Especificación de aplicación de usuario final

En esta etapa se identifican los roles o perfiles de usuarios para los diferentes tipos de aplicaciones necesarias en base al alcance de los perfiles detectados (gerencial, analista del negocio, vendedor, etc.). No todos los usuarios del Data mart necesitan el mismo nivel de análisis.

F. Selección e instalación de productos

Utilizando el diseño de arquitectura técnica como marco es necesario evaluar y seleccionar los componentes específicos de la arquitectura, como la plataforma de hardware, el motor de base de datos, la herramienta de ETL, las herramientas de acceso, etc. Una vez evaluados y seleccionados los componentes determinados se procede con la instalación y prueba de los mismos en un ambiente integrado de Data mart.

G. Diseño físico

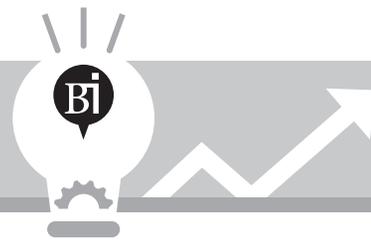
Se focaliza sobre la selección de las estructuras necesarias para soportar el diseño lógico. Un elemento principal de este proceso es la definición de estándares del entorno de la base de datos. La indexación y las estrategias de la partición se determinan también en esta etapa.

H. Diseño y desarrollo:

Esta etapa es típicamente la más subestimada de las tareas en un proyecto de Data mart, aquí se realiza la extracción, la transformación y la carga (ETL process). Es decir obtener los datos que permitirán efectuar la carga del Modelo Físico diseñado. Asimismo, en el proceso de transformación se convierten o recodifican los datos a fin de poder efectuar la carga efectiva del Modelo Físico. Por otra parte, los procesos de carga de datos son los procesos requeridos para poblar el Data mart.

I. Desarrollo de aplicación de usuario final

Después de la especificación de las aplicaciones para usuarios finales, el desarrollo involucra configuraciones de los metadatos y construcción de reportes específicos. Los usuarios acceden al DM por medio de herramientas de productividad basadas en GUI (Graphical User Interface). De hecho existen multitud de estas herramientas para proveer a los usuarios.



J. Implementación

La implementación representa la convergencia de la tecnología de solución. Hay varios factores extras que aseguran el correcto funcionamiento, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación y las estrategias de feedback. Todas estas tareas deben tenerse en cuenta antes de que cualquier usuario pueda tener acceso al DM.

I. MODELO PROPUESTO DE SOLUCIÓN

Como se mencionó inicialmente, la empresa Figueri SRL no cuenta con un sistema de inteligencia de negocios y pronóstico de ventas, y que es indispensable para la empresa tomar decisiones óptimas y a menor tiempo posible. Lo que se pretende desarrollar es una solución de inteligencia de negocios que ayude a la toma de decisiones al nivel estratégico de Figueri (FSRL). Proveen un pronóstico de ventas utilizando un modelo de pronóstico de series de tiempo, con el objetivo de determinar la proyección del comportamiento de las ventas con un mínimo porcentaje de error. Se utilizó la metodología Kimball. Esta metodología se fundamenta en el ciclo de vida dimensional del negocio, usada para el desarrollo e implementación del Data mart, un repositorio de información orientado al área específica de ventas. El desarrollo se realizó bajo el marco de un sistema que apoya a la evaluación y comparación de alternativas de solución (DSS).

Los resultados muestran información sobre la tendencia de las ventas, el porcentaje de participación de sus clientes en el total de ventas, el detalle de las ventas en meses y años, es decir un sistema que maneje indicadores y que incluya alertas de negocio.

El Data mart no es capaz de realizar un análisis predictivo por sí mismo, es decir pronósticos de ventas que es exactamente lo que necesitamos para la toma de decisiones. Para ello nos basaremos en un modelo de pronóstico de series de tiempo que tiene la función de clasificar, agrupar y predecir en base a la información (historial) que proporciona el Data mart.

En el desarrollo del modelo de pronóstico de serie de tiempo, se está utilizando el modelo au-

torregresivo, y se incluye en el reporte de la demanda de productos por ruta y semanas de la distribuidora de la empresa FSRL.

El modelo autorregresivo de orden está dado por:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p + \phi_{t-p} + \varepsilon_t \dots (1)$$

Expresado en términos del operador de retardos,

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) X_t = \varepsilon_t$$

$$\phi_1(L) X_t = \varepsilon_t$$

Donde ε_t es un proceso de ruido blanco y $\phi_1, \phi_2, \phi_3, \dots, \phi_p$ son los parámetros del modelo.

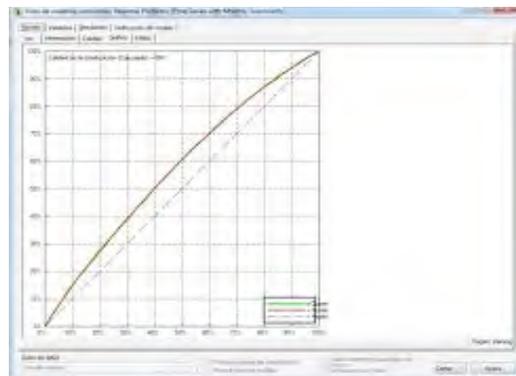


Figura 1: Pronóstico de una Variable

II. CONSTRUCCIÓN DE LA SOLUCIÓN

El sistema de inteligencias de negocios consta de la elaboración de un Data mart, que tiene cuatro dimensiones y una tabla de hechos, estos son cargados usando Microsoft SQL Server Integration Services, el paquete construido se muestra en la figura 3.

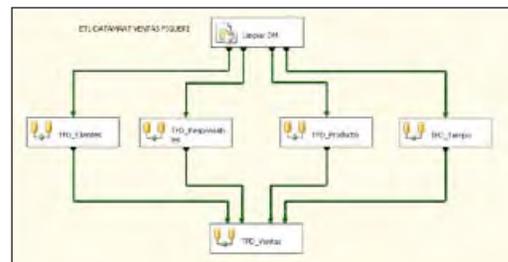
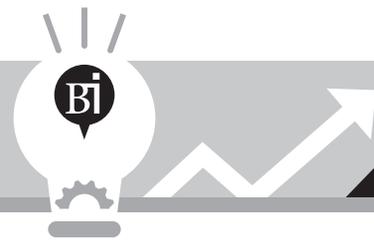


Figura 2. Paquete de ETL

Los cubos de la solución son ventas por clientes y por productos, fueron elaborados en Microstrategy Desktop. Los reportes son informaciones tomadas del Data Mart y son presentadas en la



plataforma web de Microstrategy, como se muestra en las figuras 4 y 5.

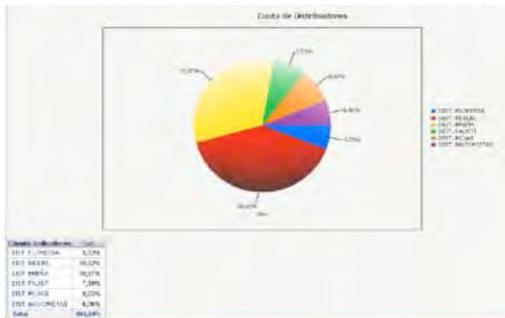


Figura 3: Cuota de Distribuidor

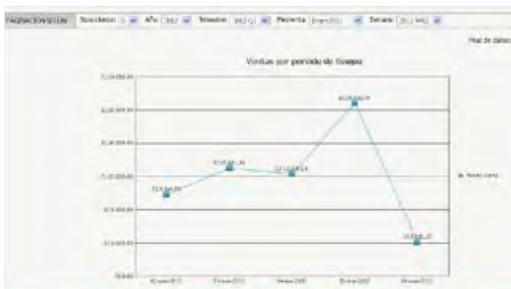


Figura 4: Tendencia de Ventas

Para el pronóstico de ventas, que será el trabajo a futuro, se desarrollará con el software SPSS Clementine 12.0. Se codificará el modelo de pronóstico de series de tiempo utilizando las capas y las variables necesarias, así como información estructurada. Podrá profundizar y predecir soluciones para poder resolver mejor los problemas de negocio que suponen un reto, como el pronóstico de las devoluciones y productos malogrados.

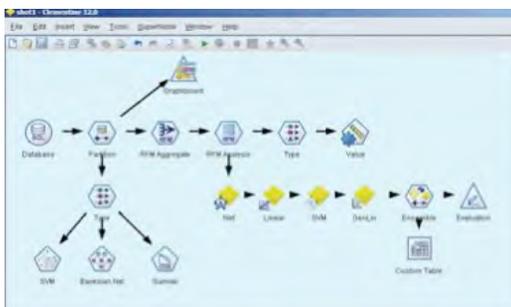


Figura 5: Clementine

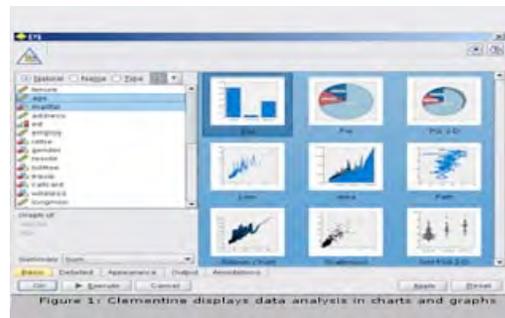


Figura 6: Gráficos de Pronósticos

I. VALIDACIÓN

El proyecto se evaluará en un periodo determinado por la empresa. Los resultados se verificarán detalladamente con los datos de la empresa.

Se utilizó la prueba estadística de Ljung-Box, para determinar la validez del modelo de pronóstico de series de tiempo.

$$LB = n(n + 2) \sum_{k=1}^m \left(\frac{\hat{\rho}_k^2}{n - k} \right) \sim X^2_{(m)}$$

Donde n tamaño de la muestra, m longitud del rezago.

Ho: Las autocorrelaciones son independientes.

Ha: Las autocorrelaciones no son independientes.

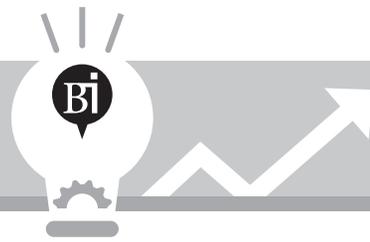
I. CONCLUSIONES

Al utilizar el sistema de soporte a la toma de decisiones, la empresa tuvo una disminución en el porcentaje de devolución, y en el porcentaje de rotación de productos.

Los pronósticos del sistema ayudaron en la toma de decisiones a futuro, generando mayores ventas, y nuevas estrategias de marketing.

Las series de tiempo agregan mayor confianza al considerar la estacionalidad, y los ciclos de las ventas de la empresa FSRL, ya que los productos de esta empresa tienen estacionalidad y ciclos de demanda fluctuantes.

Después de haber aplicado la prueba estadística de Ljung-Box, si el Q calculada excede el valor



Q crítico de la tabla ji cuadrada al nivel de significancia seleccionado, no se acepta la hipótesis nula de que todos los coeficientes de autocorrelación son iguales a cero; por lo menos algunos de ellos deben ser diferentes de cero.

Se recomienda utilizar este tipo de sistema en empresas medianas y grandes, porque agrega una opción para tomar decisiones respecto a las clásicas soluciones de inteligencia de negocios.

FIGURAS

Figura 2: Pronóstico de una variable.....	7
Figura 3. Paquete de ETL.....	7
Figura 4: Cuota de Distribuidor.....	7
Figura 5: Tendencia de Ventas.....	7
Figura 6: Clementine	7
Figura 7: Gráficos de Pronósticos.....	8

REFERENCIAS

- [1] V. Bures, T. Otcenaskovs, and V. Jasikova, "The evaluation of external data resource for business intelligence applications: the example of the Czech Republic", *Journal of Systems Integration*, 2012.
- [2] M. Bahrami, S. M. Arabzad, and M. Ghorbani, "Innovation In Market Management By Utilizing Business Intelligence: Introducing Proposed Framework," *Procedia - Social and Behavioral Sciences*, vol. 41, pp. 160-167, 2012.
- [3] R. Sanchez, *Business Intelligence (BI) - To BI or not to BI*. Argentina: El Cid Editor | apuntes, 2009.
- [4] F. Zada, S. K. Guirguis, and A. H. Sedky, "Development of a Dynamic Model for Data-Driven DSS", *Journal of Emerging Trends in Computing and Information Sciences*, 2012.
- [5] N. F. o. E. Statistics, *Forum Guide to Decision Support Systems: A Resource for Educators (NFES 2006-807)*. Washington, DC: National Center for Education Statistics, 2006.
- [6] M. M. Azad, M. B. Amin, and M. Alauddin, "Executive Information System", *IJCSNS International Journal of Computer Science and Network Security*, 2012.
- [7] I. Lungu, "Executive Information Systems", *Revista Informatica Economica*, 2005.
- [8] N. L. Parracia, *Datawarehousing*. Argentina: El Cid Editor | apuntes, 2009.
- [9] S. H. Ha and S. C. Park, "Application of data mining tools to hotel data mart on the Intranet for database marketing", *Expert Systems with Applications*, 1998.
- [10] P. Das and S. Chaudhury, "Prediction of retail sales of footwear using feedforward and recurrent neural networks", *Neural Computing and Applications*, vol. 16, pp. 491-502, 2006.
- [11] H. Cruse, *Neural Networks as Cybernetic Systems*. Germany: Brains, Minds & Media, 2006.
- [12] B. Y. Shih and Y. S. Chung, "The Development of a CFM Hybrid Artificial sale Forecasting Model", *International Journal of Electronic Business Management*, 2008.
- [13] R. Kimball, *The Data Warehouse Lifecycle Toolkit*. New York: Wiley, 2008.