

CAPÍTULO I

Introducción

El objetivo de este capítulo es mostrar de manera general el contexto en el que se desenvuelve esta tesis, los conceptos sobre los que se fundamenta, los objetivos y la organización del documento. El capítulo se encuentra organizado de la siguiente manera: la sección 1.1 corresponde contexto de la tesis, la sección 1.2 corresponde a los conceptos fundamentales, en la sección 1.3 se describe la problemática, en la sección 1.4 se explica la motivación, las secciones 1.5 y 1.6 corresponden a los objetivos general y específicos respectivamente, en la sección 1.7 se describe la metodología que se siguió en la investigación y por último en la sección 1.8 se define la organización del documento.

1.1 Las PYMEs y la toma de decisiones

Las PyMEs (Pequeñas y Medianas Empresas), son organizaciones que se caracterizan por tener un número reducido de empleados y una facturación media o baja. En el censo económico del INEGI (Instituto Nacional de Estadística Geografía e Informática) del 2004, se establece que existen en México alrededor de 2,726,568 PyMEs, las cuales representan aproximadamente el 94% del total de las empresas mexicanas [Diario Oficial de la Federación, 2004].

La Secretaría de Comercio y Fomento Industrial ha clasificado a las PyMEs por sector: manufactura, comercio y servicio, y a su vez por el número de empleados con los que cuentan, como se muestra en la tabla 1.1.

Tamaño	Clasificación de sectores por número de Empleados		
	Industria	Comercio	Servicio
Pequeña Empresa	31-100	6-20	21-50
Mediana Empresa	101-500	21-100	51-100

Tabla 1.1 Estratificación de PyMEs [Diario Oficial de la Federación, 2004]

El número de empresas que conforman el sector manufacturero es de 344,118 el sector comercial cuenta con 1,443,878 y el sector servicio cuenta con 938,572 empresas. Como se puede observar, el sector comercial es el que abarca el mayor número de organizaciones, 52% del total de las PyME, y sigue en aumento [Diario Oficial de la Federación, 2004].

Con estos resultados, se puede determinar que las PyME comercializadoras, son las que conforman una parte importante de las empresas productivas de México.

Definidos el concepto de PyME y su clasificación en México, en seguida conoceremos el panorama general de la situación empresarial y la importancia de la toma de decisiones de manera oportuna.

“El crecimiento de las PyMEs ocurre actualmente entre una fuerte competencia, contracción económica, clientela más inteligente, reducción de márgenes de utilidad y constantes innovaciones tecnológicas.”[González, 2006]. De ahí la importancia de que las PyMEs tomen decisiones acertadas en el momento que se requieran y antes que su competencia lo haga.

La información es un factor crítico en los negocios. Por esto, las empresas requieren de sistemas y soluciones que permitan la exploración de la información, donde sus propios datos sean procesados para apoyar la toma de decisiones estratégicas, justificadas con información esencial para ellas.

Los ICRs (Indicadores Clave de Rendimiento), propuestos por Ronald Daniel y Jack F. Rockart, también conocidos como indicadores clave de resultados o de desempeño, son un conjunto de medidas cuantificables utilizadas por las empresas para evaluar su desempeño en términos del conocimiento de sus estrategias y metas operacionales. [Baker, 2006] El monitoreo y predicción de estos indicadores permite a las organizaciones tomar decisiones con base en información consultada en tiempo real.

El mantenerse al tanto de la situación de la empresa, les permite establecer o mejorar estrategias que generen ventajas ante su competencia, con mayor seguridad.

En seguida veremos los conceptos sobre los que se basa esta tesis, los cuales apoyan a las empresas para tomar mejores decisiones en sus procesos de negocio.

1.2 Antecedentes

La Inteligencia Empresarial (*Business Intelligence*) es, para el área de tecnologías de la información, “una arquitectura y una colección de aplicaciones operacionales integradas, así como de bases de datos que dan soporte a la toma de decisiones y que proveen a las empresas un fácil acceso a sus propios datos” [Moss, et.al., 2003].

La inteligencia empresarial se compone de varias tecnologías, que son:

- Almacenes de datos (*Datawarehouses*) que son: “colecciones de datos diseñados para dar soporte a la toma de decisiones, las cuales son integradas, no volátiles, variables en el tiempo y están orientadas a la información” [Inmon, 2005].

- OLAP (*On-Line Analytical Processing*) que permite “consultas en línea a estructuras multidimensionales, las cuales proporcionan facilidades para manejar y transformar los datos” [Hernández, 2005].
- Minería de datos (*Data Mining*) que “persigue el descubrimiento automático del conocimiento contenido en la información almacenada de modo ordenado en grandes bases de datos” [Pérez, et.al., 2006].

Estas 3 tecnologías, permiten fortalecer las capacidades de las empresas para identificar oportunidades, detectar riesgos y competir más eficaz y eficientemente.

Existen diversos procesos de negocios y marketing, a los que la inteligencia empresarial puede apoyar, por ejemplo: planeación de recursos empresariales (ERP), administración de ventas, administración de la cadena de suministro (SCM), reingeniería de procesos de negocios (BRP) y administración de la relaciones con el cliente (CRM).

El servicio diferenciado o CRM, es:

El conjunto de estrategias empresariales, marketing, comunicación e infraestructuras tecnológicas, diseñadas con el objetivo de construir una relación duradera con los clientes, identificando, comprendiendo y satisfaciendo sus necesidades [Quevedo, et.al., 2006].

Existen varias técnicas de CRM que permiten fortalecer el contacto con los clientes y construir una relación a largo plazo con ellos. Las principales técnicas son: *target marketing* (TM), *risk analysis-credit scoring* (RACS), *clustering analysis* (CA), y *market basket analysis* (MBA).

La técnica de MBA o análisis de la canasta de mercado, permite dar a conocer el comportamiento de compra de los clientes, determinando grupos o clusters de productos que tiendan a presentarse juntos en una transacción o compra realizada por el cliente.

1.3 Descripción del problema

Hoy en día las técnicas de inteligencia empresarial no se aplican comúnmente a las PyMEs, por que se piensa que es un proceso muy costoso, no solo para su creación, sino también para su mantenimiento y que se necesita de personal experto y capacitado para poder lograr buenos resultados [Harjinder, 1996] [Mallach, 2000] [Shaw, 2001].

Así mismo, las PyMEs por considerarse a ellas mismas pequeñas compañías, no creen que tengan la posibilidad y necesidad de beneficiarse de estas técnicas que les pueden ayudar a tomar decisiones para mejorar sus procesos de negocio.

Los vendedores de software de inteligencia empresarial, mencionan que la tecnología que ellos tienen puede proveer la mejora de procesos que buscan las organizaciones. Sin embargo, estas herramientas son muy costosas y complejas porque tratan de satisfacer las necesidades de las grandes empresas. Generalmente se venden como módulos separados, lo que complica de manera significativa su uso y comprensión, y eleva los costos.

La figura 1.1 muestra la arquitectura de un sistema de inteligencia empresarial para una empresa grande. Como podemos observar se tiene un sistema robusto, que si bien funciona para las grandes empresas, se encuentra excedido para una PyME.

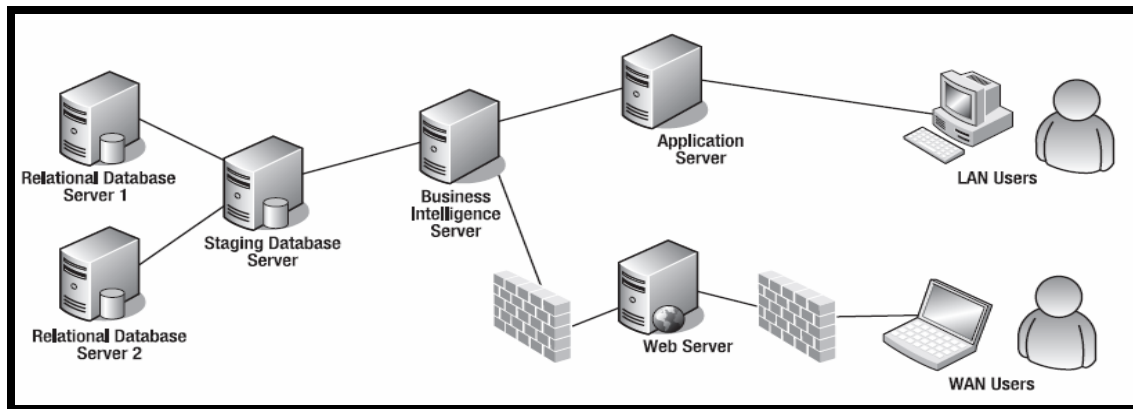


Figura 1.1 Arquitectura de un sistema de inteligencia empresarial [Langit, 2007]

Este tipo de sistemas tienen altos costos de adquisición y mantenimiento, como se muestra en la tabla 1.2 con algunos de los sistemas de inteligencia empresarial más populares utilizados en las grandes empresas.

Sistemas	Oracle BI 10g	Microsoft SQL Server 2005 (SSIS, SSAS, Data Mining)	Cognos 8 BI
Precio	\$40,00 usd	\$24,499 usd	\$220,000 usd

Tabla 1.2 Costos de sistemas de inteligencia empresarial [Microsoft, 2007] [Aware, 2007]

El sistema de Cognos tiene un precio elevado porque incluye algunas de las siguientes características: metadatos centralizados, interfaces de usuario independientes de las estructuras de los datos, las funcionalidades de OLAP para fuentes de datos relacionales y una interfaz basada en *browser*.

Debido a la complejidad, los sistemas requieren de personal experto en el área para administrarlos, por ejemplo para la suite de Cognos se sugiere el siguiente grupo: 15 consumidores BI, 15 responsables de negocios, 15 analistas de negocio, 10 administradores de negocio, 5 expertos en BI y 2 administradores de BI [CRM, 2005]. En el caso de Oracle BI 10g

se sugiere los grupos de personas que vemos en la figura 1.2. Por estas razones, sistemas como éstos, no se ajustan a las necesidades tanto económicas como tecnológicas de las PyMEs.

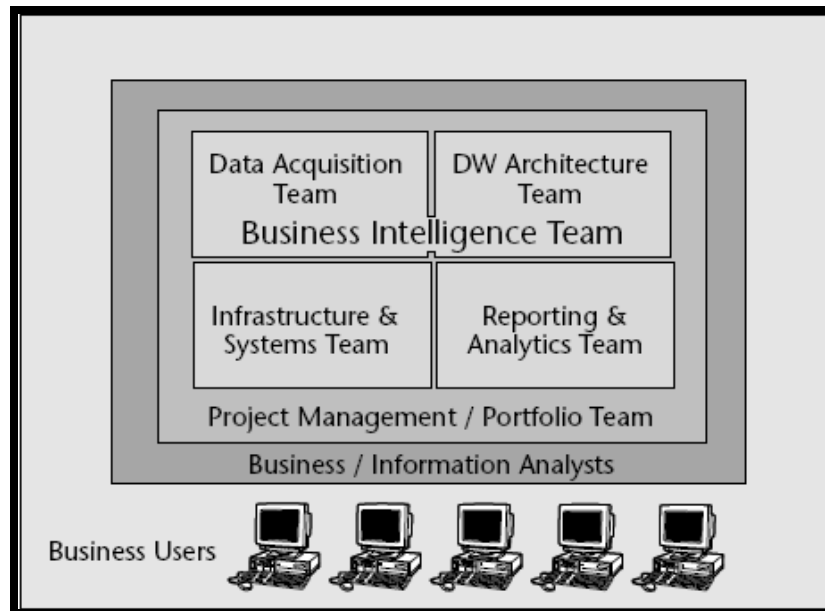


Figura 1.2 Staff propuesto por Oracle BI para un sistema de inteligencia empresarial [Stackowiak, 2007]

1.4 Motivación

Las PyMEs tienen las mismas necesidades de análisis de información que las empresas de gran tamaño, sin embargo tienen muchas más desventajas y limitaciones que estas últimas.

Las limitaciones económicas son el principal factor, por el que las pequeñas y medianas empresas no adquieren la tecnología que requieren. Además, los sistemas existentes en el mercado, no ofrecen los resultados que prometen, se concentran más en la venta de los mismos que en resolver los problemas de los usuarios.

Otro motivo son los numerosos beneficios que las PyMEs pueden obtener, de los avances de la tecnología que ahora permiten procesar datos a la velocidad del pensamiento [González, 2006], por ejemplo:

- Liberar a ciertos empleados de tiempo operativo para dedicarlo a realizar análisis estratégicos.
- Tener información suficiente que justifique la toma de decisiones estratégicas en un momento determinado.
- Aprovechar mejor de sus fuentes de información.
- Fortalecer de las capacidades analíticas y de planificación.
- Mejorar el entendimiento de las necesidades de los clientes

La creación de sistemas de tecnologías de información accesibles para las PyME contribuirá a que tengan un crecimiento más rápido y estén mejor preparadas para el ambiente competitivo en el que se desempeñan.

1.5 Objetivo general

El objetivo general de esta tesis es el diseño e implementación de un prototipo que integre las tecnologías de inteligencia empresarial: *datawarehousing*, OLAP y *data mining*, ofreciendo un ambiente integral y factible para las necesidades de una PyME comercializadora.

El prototipo está enfocado a apoyar la toma de decisiones de las PyMEs comercializadoras, en el área de *Customer Relationship Management* (CRM) utilizando en particular, la técnica de *Market Basket Analysis* (MBA).

Las herramientas que se utilizaron para el desarrollo del prototipo son de código abierto, esto con el fin de que sea económicamente accesible para una PyME.

La tesis es de tipo aplicativo e integra distintas herramientas de software ya existentes, que cumplen con diversos estándares entre ellos el estándar de *data mining*: CRISP-DM (Cross Industry Standard Process for *Data mining*) [CRISP-DM, 2006].

1.6 Objetivos específicos

Los objetivos específicos propuestos para esta tesis fueron:

1. Realizar un prototipo de un sistema de inteligencia empresarial, integrando las técnicas de *data warehousing*, OLAP y *data mining*, que responda a las necesidades de las PyMEs de tipo comercial.
2. Aprender y conocer de manera general las funcionalidades de algunas de las herramientas existentes en el mercado para cada una de las técnicas mencionadas anteriormente.
3. Evaluar y comparar de manera general las herramientas investigadas y elegir una(s) que sean útiles para la problemática que den soporte a las tecnologías de inteligencia empresarial.

4. Proponer un modelo multidimensional para soporte a la toma de decisiones en el área de CRM
5. Identificar un modelo de comportamiento de compra del cliente, que sea representativo para los análisis de la información.
6. Diseñar e implementar un ambiente integral que facilite el uso de las herramientas seleccionadas y la implementación del procedimiento de inteligencia empresarial en las PyMEs comercializadoras.

1.7 Metodología de investigación

La metodología que se siguió para la realización de esta tesis, organizó las actividades en tres bloques que corresponden a los temas que fundamentan la tesis: *data warehousing*, OLAP y *data mining*. De manera general, se definen actividades organizadas en cuatro fases: lectura, especificación, implementación y validación:

- **Lectura:** en esta fase se recopila y selecciona la información pertinente que servirá para definir y delimitar el escenario de prueba del sistema, así como para conocer los trabajos existentes del área.
- **Especificación:** en esta fase se hace el diseño del sistema que resuelve un aspecto de la inteligencia empresarial.
- **Implementación:** en esta fase se escogen las herramientas adecuadas para la implementación del diseño que se propuso en la fase de especificación
- **Validación:** se prueba la implementación en el contexto del escenario de las PyMEs.

Las fases de lectura y validación son fases generales para los tres bloques mientras que las de especificación e implementación son iterativas para cada uno de ellos.

1.7.1 Escenario

Debido a que el objetivo principal era construir un prototipo que integre las técnicas de inteligencia empresarial, que sea adecuado para una PyME, se buscó y recopiló información necesaria para la aplicación de las técnicas mencionadas anteriormente: *datawarehousing*, OLAP y *data mining* en una PyME comercial abarrotera.

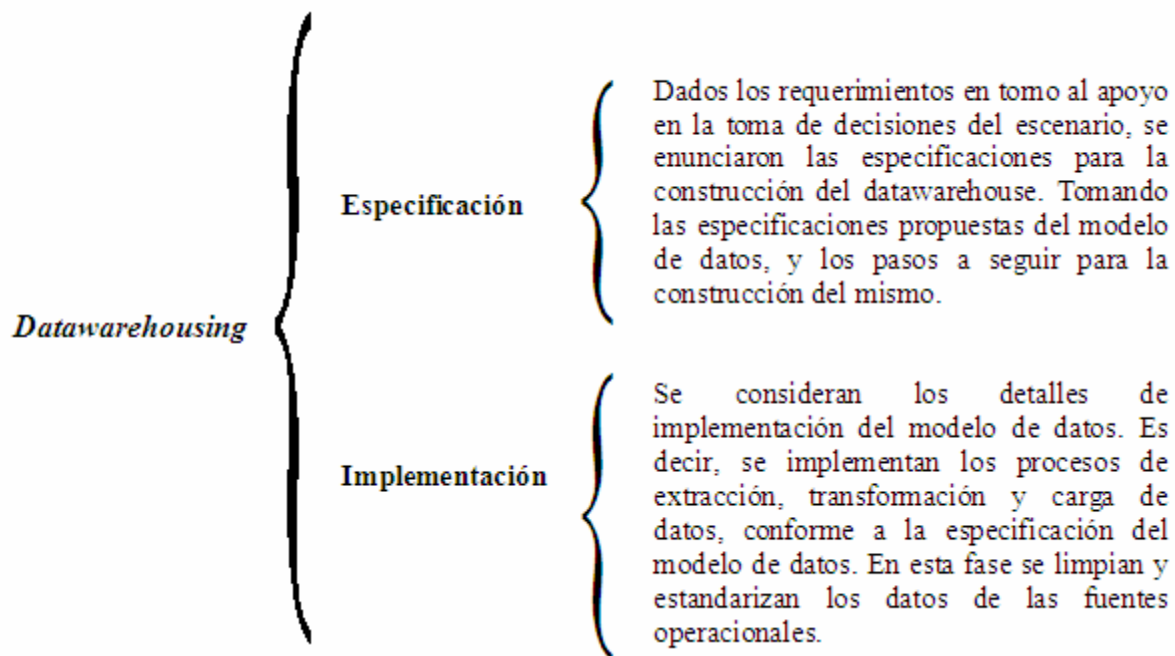
El proceso de inteligencia empresarial estuvo dirigido al área de CRM, con el fin de fortalecer las relaciones con clientes, haciendo uso específicamente de la técnica de MBA.

1.7.2 Lectura

Se busca y recopila información necesaria para comprender los conceptos fundamentales para la construcción de un *datawarehouse*, la construcción de cubos o hipercubos para realizar análisis multidimensional y la aplicación de los algoritmos de *data mining* a los datos almacenados, así como para conocer los trabajos similares, realizados hasta la fecha, respecto al tema. La información fue organizada en términos de análisis, diseño, construcción e implementación y pruebas.

1.7.3 Datawarehousing

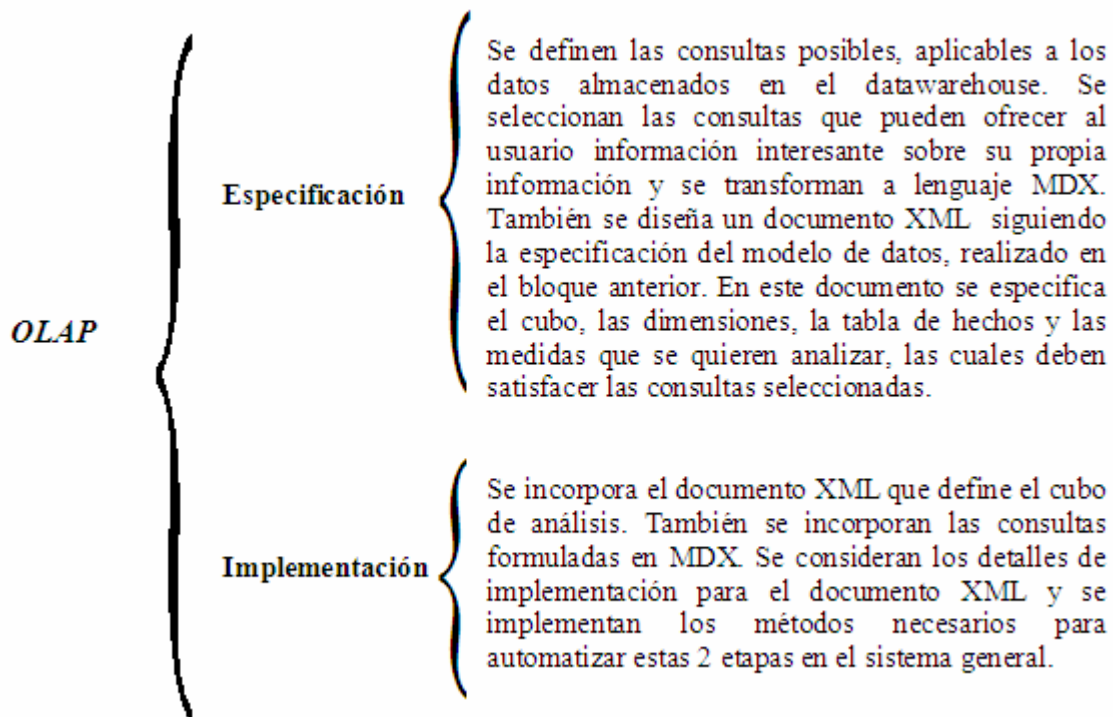
Este bloque comprende todo lo relacionado con el diseño e implementación de un *datawarehouse* para PyMEs. En seguida se detallan las fases para este bloque:



Como resultado de este bloque tenemos el conocimiento teórico necesario para implementar un *datawarehouse*, así como los pasos a seguir para la implementación física de un *datawarehouse* con datos reales de una PyME comercializadora.

1.7.4 OLAP

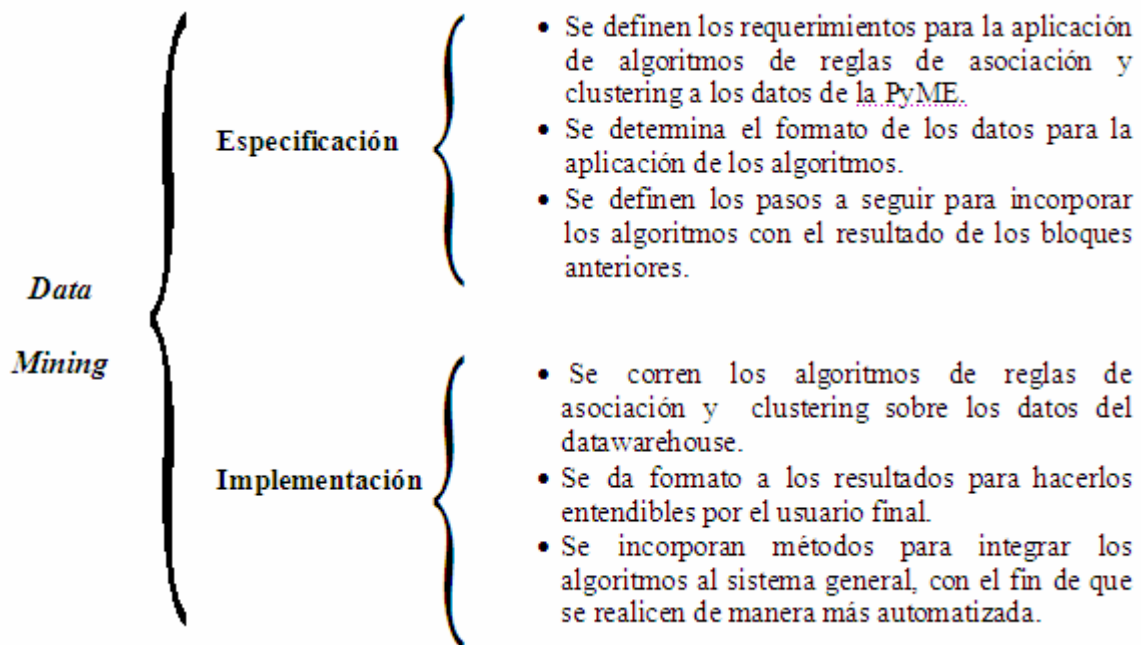
Este bloque abarca todo lo relacionado con la generación de cubos o hipercubos, que permiten explotar la información almacenada en el *datawarehouse* de acuerdo a las necesidades de consulta del usuario final, de una manera entendible y amigable. Las fases detalladas de este bloque son:



Como resultado de este bloque se obtuvo conocimiento teórico respecto a la elaboración de cubos de OLAP para el análisis de información y su implementación física mediante documentos XML que permiten manipular los datos de la PyME almacenados en el *datawarehouse*. Al final de este bloque se encuentran integradas las técnicas de *datawarehousing* y OLAP.

1.7.5 Data Mining

Este bloque incluye todo lo relacionado con *data mining*, la técnica que permitirá obtener reglas de asociación y describir comportamiento de compra de los clientes de la PyME comercializadora. Las fases detalladas del bloque son las siguientes:



Como resultado de este bloque se obtuvieron reglas de asociación y grupos que van a permitir realizar predicciones sobre el total de los datos reales de la PyME. Así mismo, se obtuvo el conocimiento teórico para la aplicación de algoritmos de *data mining* y se incorpora este bloque con el resultado de los 2 bloques anteriores.

1.7.6 Validación

Se definió un escenario de simulación, que ayude a probar los procesos de ETL, las consultas predeterminadas sobre el modelo de datos multidimensional y la incorporación del documento XML sobre el que se realizan las consultas, así como la aplicación de los algoritmos de *data mining* de reglas de asociación y clustering sobre los datos del *datawarehouse*.

El escenario de simulación consistió en tomar datos reales de una PyME comercial abarrotera sobre los que se aplicaron las actividades mencionadas anteriormente.

Con esta metodología se obtuvo como resultado final un conocimiento teórico- práctico, sobre la implementación de un sistema de inteligencia empresarial en una PyME comercializadora abarrotera.

Así mismo se verificó la factibilidad de implementar las técnicas de inteligencia empresarial en otras PyMEs comercializadoras, aplicando las actividades mencionadas en la metodología a los datos de otra PyME comercializadora, y se obtuvo el sistema integrado desarrollado a lo largo de los bloques mencionados anteriormente.

1.8 Organización del documento

Este documento está organizado en 6 capítulos. Enseguida veremos una descripción general de cada uno de ellos:

Capítulo 1 – Introducción

En este capítulo se dan a conocer el contexto del sistema: las PyMEs y la toma de decisiones. Se menciona de manera general, los conceptos de inteligencia empresarial y la descripción del problema que se aborda, el objetivo general y los específicos que se pretenden alcanzar, la metodología empleada en la investigación, así como la organización del documento.

Capítulo 2 – Inteligencia Empresarial

En este capítulo se mencionan más detalladamente los conceptos fundamentales de la inteligencia empresarial: *datawarehousing*, OLAP y *data mining*, que se utilizaron para la

construcción del sistema. También se muestran los conceptos del área de negocios: CRM y MBA, así como algunos ejemplos para ilustrarlos.

Capítulo 3 – Herramientas y software de inteligencia empresarial

En este capítulo se describe una clasificación de las herramientas y software que existen en el mercado relacionadas con el área de inteligencia empresarial, tanto comerciales como de código abierto.

Capítulo 4 – Caso de estudio

En este capítulo se describe el contexto aplicativo de la tesis: “Abarrotera Coscomatepec”, su situación actual con respecto a la toma de decisiones en el área de CRM y un análisis de sus requerimientos de tecnologías de la información. También se muestra el análisis y diseño realizados para determinar la arquitectura y estructura del sistema. Se pueden observar diagramas y modelos que se emplearon para una mayor comprensión del mismo.

Capítulo 5 – Implementación

En este capítulo se muestra la implementación del ambiente integrado en una interfaz gráfica. También se da una explicación de su funcionamiento y las herramientas utilizadas para la construcción del prototipo: Kettle, Weka y Mondrian.

Capítulo 6 – Estado actual y perspectivas

En éste capítulo se mencionan los resultados obtenidos de la tesis, la metodología de aplicación de sistemas de inteligencia empresarial en las PyMEs, el estado actual de la aplicación,

las pruebas, evaluaciones y validaciones que se hicieron. Por último, se describen las perspectivas a futuro y mejoras al sistema.

Dado que el lector puede ser una persona con conocimientos en el tema, la figura 1.3 muestra la secuencia sugerida de lectura de los capítulos de esta tesis.

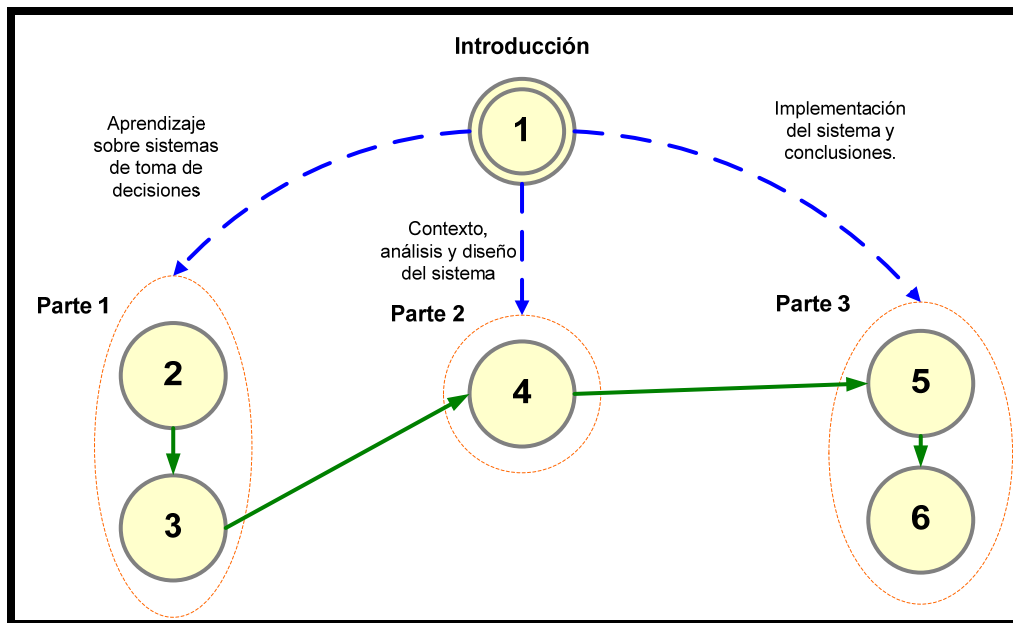


Figura 1.3 Secuencia sugerida de lectura de capítulos