

## 1. Datos Generales de la asignatura

<b>Nombre de la asignatura:</b>	<b>Inteligencia de Negocios.</b>
<b>Clave de la asignatura:</b>	<b>IFF-1016</b>
<b>SATCA<sup>1</sup>:</b>	<b>3 - 2 - 5</b>
<b>Carrera:</b>	<b>Ingeniería en Informática.</b>

## 2. Presentación

### Caracterización de la asignatura

La Inteligencia de Negocios es el proceso de integración y tratamiento de los datos para convertirlos en información que permita apoyar a los tomadores de decisiones en la organización. Aporta al perfil del Ingeniero en Informática en las siguientes competencias:

- Aplica conocimientos científicos y tecnológicos en el área informática para la solución de problemas con un enfoque multidisciplinario.
- Aplica herramientas computacionales actuales y emergentes para optimizar los procesos en las organizaciones.
- Diseña e implementa Bases de Datos para el almacenamiento, recuperación, distribución, visualización y manejo de la información en las organizaciones.
- Realiza consultorías relacionadas con la función informática para la mejora continua de la organización.
- Se desempeña profesionalmente con ética, respetando el marco legal, la pluralidad y la conservación del medio ambiente.
- Participa y dirige grupos de trabajo interdisciplinarios, para el desarrollo de proyectos que requieran soluciones innovadores basadas en tecnologías y sistemas de información.

La Inteligencia de Negocios aporta, al perfil del ingeniero, la capacidad para comprender desde la extracción de los datos de sistemas existentes hasta la explotación de la información por herramientas de análisis de datos.

Esta asignatura sirve para el descubrimiento de conocimiento en bases de datos que posee una empresa. Permite la creación de almacenes de datos (data warehouse) hasta la utilización de herramientas de minería de datos para el soporte en la toma de decisiones. Proporciona al estudiante de ingeniería un conjunto de mecanismos para el análisis dirigido por los datos, los cuales permiten moverse a través de los almacenes de datos para encontrar las tendencias, patrones y correlaciones que pueden guiar la toma de decisiones estratégicas.

Está diseñada para el logro de las siguientes competencias específicas dirigidas a la aprehensión de los dominios: datawarehouse o datamart, almacenes de datos multidimensionales, herramientas de visualización, pivoteo y consultas en línea, , minería de datos, OLAP.

<sup>1</sup> Sistema de Asignación y Transferencia de Créditos Académicos

La intención de esta asignatura es que el egresado construya soluciones de inteligencias de negocios que apoyen la toma de decisiones en una empresa. Se pretende que conozca y aplique las tecnologías emergentes de bases de datos, así como, las metodologías y tecnologías existentes para el desarrollo de las soluciones. El reto es formar egresados con la capacidad de proponer a las empresas soluciones que le permitan un análisis estratégico que contribuya a la mejora de su productividad y rentabilidad, para lograr una ventaja estratégica en el mundo altamente competitivo.

### **Intención didáctica**

La asignatura pretende proporcionar al estudiante los conceptos esenciales de la Inteligencia de Negocios. Se organiza el contenido en cuatro temas.

En el primer tema se estudian los conceptos básicos de la Inteligencia de Negocios ofreciendo y proporcionando una visión integral de la importancia que tiene en el entorno actual y cómo utilizarla como herramienta estratégica para alcanzar los objetivos de la organización. Entre los conceptos básicos que se abordan están la definición de la Inteligencia de Negocios, la descripción de la arquitectura general describiendo sus elementos, definición de las propiedades como amplitud, profundidad, calidad de los datos, almacenamiento inteligente y escalable, análisis avanzado por mencionar algunas. Se enfatizará en este tema que la Inteligencia de Negocios se compone de todas las actividades relacionadas a la organización y entrega de información así como el análisis del negocio. Esto incluye Minería de Datos, Administración del Conocimiento, Aplicaciones Analíticas, Sistemas de Reportes y principalmente Data Warehousing. Así también se abordan en este tema las principales herramientas para la Inteligencia de Negocios tales como Oracle, DB2, Microsoft SQL Server, Wired for OLAP, entre otras.

En el segundo tema se aborda el uso de Bases de Datos para la Toma de Decisiones. Primeramente, se discute y analiza el uso de base de datos multidimensionales vista como una colección de datos sistematizados, integrados, variables en el tiempo para dar soporte al proceso de toma de decisiones. Este proceso aglutina datos de fuentes heterogéneas e involucra esfuerzos de toda la compañía para que el soporte a decisiones actúe en todos los niveles de la empresa. En las bases de datos multidimensionales, los datos antes de ser almacenados son filtrados, normalizados, reorganizados, resumidos para constituir una base de datos confiable. Después se abordan los sistemas de adquisición de conocimiento que procesan automáticamente grandes cantidades de datos para encontrar conocimiento útil en ellos, de esta manera permitirá al usuario el uso de esta información valiosa para la toma de decisiones. Esto se lleva a cabo mediante un proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles a partir de los datos, teniendo como objetivo encontrar conocimiento útil relevante y nuevo sobre un fenómeno o actividad, presentando los resultados de manera visual.

En el tercer tema se estudian los componentes del Entorno de Inteligencia de Negocios como la minería de datos, data marts, el procesamiento analítico On-line y las herramientas de análisis de datos. Se estudia que la minería de datos es una técnica que utiliza herramientas de software, generalmente orientadas para los usuarios que no saben exactamente lo que están investigando, mas procura identificar determinados patrones o tendencias. Es un proceso que separa grandes cantidades de datos de forma que identifica relaciones entre estos. Toda la información escondida relacionada al comportamiento de los clientes es mapeada y enfatizada. La minería de datos provee cinco tipos de información:

(1) Asociación: Ocurrencias están ligadas a un evento. Por ejemplo: en un supermercado, cuando se compra atún, se compra también mayonesa. (2) Secuencia: Eventos ligados a lo largo del tiempo. Por ejemplo si una casa es comprada, entonces 45% de las veces una cocina será comprada dentro de un mes. (3) Clasificación: Encontrar características de los clientes y determinar cuáles tipos de promociones serán más eficientes. (4) Conjuntos: descubrir diferentes agrupamientos de datos. Utilizando, por ejemplo para descubrir defectos en la fabricación de productos (5) Previsión: Si un cliente renueva el contrato, por ejemplo. Así también, se estudia el procesamiento analítico On-line que comprende un conjunto de herramientas que posibilita efectuar la exploración de los datos contenidos en un data warehouse. Se enfatiza el uso de Análisis multidimensional y se discute el proceso interactivo en que el usuario hace preguntas, recibe informaciones, verifica un dato específico y hace comparaciones. Como ejemplo de preguntas que un OLAP responde se tiene: ¿Porqué las ventas de automóviles de un determinado modelo disminuyeron en el segundo trimestre?

Finalmente, en el cuarto tema se aborda la construcción a la solución de la inteligencia de negocios. Se aborda la creación, desarrollo e integración de un proyecto con las técnicas apropiadas para la toma de decisiones en la inteligencia de negocios.

Los contenidos presentados constituyen los elementos básicos indispensables de la Inteligencia de Negocios. Para abordar estos contenidos se proponen actividades de aprendizaje que permitan al alumno conocer la herramientas que dan origen a los conceptos básicos, y a partir de ellas extender el conocimiento.

Las actividades de aprendizaje recomendadas pretenden servir de ejemplo para el desarrollo de las competencias, mencionadas más adelante en este documento, y se propone adecuarlas a la especialidad y al contexto institucional.

### 3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Evento
Instituto Tecnológico de Saltillo del 5 al 9 de octubre de 2009.	Representantes de los Institutos Tecnológicos de:  Apizaco, Cerro Azul, Chetumal, Ciudad Juárez, Ciudad Madero, Superior de Coahuila de Zaragoza, Colima, Comitancillo, Conkal, Durango, El Llano Aguascalientes, El Salto, Superior de Fresnillo, Huejutla, Superior de Lerdo, Linares, Los Mochis, Mexicali, Morelia, Oaxaca, Superior del Occidente del Estado de Hidalgo,	Reunión Nacional de Diseño e Innovación Curricular para el Desarrollo y Formación de Competencias Profesionales de las Carreras de Ingeniería en Sistemas Computacionales, Ingeniería Informática e Ingeniería en Geociencias.

	Ocotlán, Orizaba, Piedras Negras, Pinotepa, Saltillo, San Luis Potosí, Tapachula, Tijuana, Torreón, Tuxtepec, Superior de Valladolid, Valle del Guadiana, Superior de Zacapoaxtla y Zacatecas.	
Instituto Tecnológico Superior de Poza Rica del 22 al 26 de febrero de 2010.	Representantes de los Institutos Tecnológicos de:  Apizaco, Cerro Azul, Chetumal, Ciudad Juárez, Ciudad Madero, Superior de Coahuila de Zaragoza, Colima, Comitancillo, Conkal, Durango, El Llano Aguascalientes, El Salto, Superior de Fresnillo, Huejutla, Superior de Lerdo, Los Mochis, Mexicali, Morelia, Oaxaca, Superior del Occidente del Estado de Hidalgo, Ocotlán, Orizaba, Piedras Negras, Pinotepa, Saltillo, San Luis Potosí, Tapachula, Tijuana, Torreón, Tuxtepec, Superior de Valladolid, Valle del Guadiana, Superior de Zacapoaxtla y Zacatecas.	Reunión Nacional de Consolidación de los Programas en Competencias Profesionales de las Carreras de Ingeniería en Sistemas Computacionales, Ingeniería Informática e Ingeniería Petrolera del SNEST.
Instituto Tecnológico de Querétaro del 22 al 25 de octubre de 2012.	Representantes de los Institutos Tecnológicos de:  Acayucan, Campeche, Cd. Madero, Celaya, Chilpancingo, Coahuila de Zaragoza, Colima, Ecatepec, El Grullo, Iguala, Jiquilpan, Lerdo, Los Mochis, Morelia, La Región Sierra, San Andrés Tuxtla, Sur de Guanajuato, Teziutlán, Tizimín, Zacatecas y Zitácuaro.	Reunión Nacional de Seguimiento Curricular de los Programas en Competencias Profesionales de las Carreras de Ingeniería en Sistemas Computacionales, Ingeniería Informática e Ingeniería en Tecnologías de la Información y Comunicaciones.

<p>Instituto Tecnológico de Toluca, del 10 al 13 de febrero de 2014.</p>	<p>Representantes de los Institutos Tecnológicos de:  Cerro Azul, Colima, Lerdo, Toluca y Veracruz.</p>	<p>Reunión de Seguimiento Curricular de los Programas Educativos de Ingenierías, Licenciaturas y Asignaturas Comunes del SNIT.</p>
--	---	--

#### 4. Competencia(s) a desarrollar

Competencia(s) específica(s) de la asignatura
<p>Aplicar Tecnologías Emergentes de Base de Datos para construir soluciones de Inteligencia de Negocios de soporte a la Toma de Decisiones.</p>

#### 5. Competencias previas

<ul style="list-style-type: none"> <li>• Diseñar y manipular bases de datos relacionales</li> <li>• Administrar bases de datos utilizando sistemas de gestión de base de datos.</li> </ul>
--

#### 6. Temario

No.	Temas	Subtemas
1	Introducción a la Inteligencia de Negocios.	<p>1.1. Conceptos básicos.            1.1.1. Data Warehouse.            1.1.2. Data Mart.            1.1.3. Tipos de sistemas de información.            1.1.4. Variables de medición.            1.1.5. Variables de análisis.</p> <p>1.2. Componentes de la Inteligencia de Negocios.            1.2.1. Minería de Datos.            1.2.2. Administración del Conocimiento.            1.2.3. Aplicaciones Analíticas.            1.2.4. Sistemas de Reportes.            1.2.5. Multidimensionalidad.            1.2.6. Data Warehouse.</p> <p>1.3. Principales Herramientas de la Inteligencia de Negocios.</p>
2	Base de Datos para la Toma de Decisiones.	<p>2.1. Base de Datos Multidimensionales.            2.1.1. Data Warehouse            2.1.2. Data Mart            2.1.3. Sistemas OLTP</p>



		<ul style="list-style-type: none"> <li>2.1.4. Sistemas OLAP</li> <li>2.1.5. Operaciones Analíticas Básicas de los Sistemas OLAP</li> <li>2.1.6. Vista de Datos de los sistemas OLAP</li> <li>2.1.7. Modelo de Datos de los sistemas OLAP.</li> <li>2.2. Sistemas de Gestión del conocimiento.                             <ul style="list-style-type: none"> <li>2.2.1. Preparación de los Datos.</li> <li>2.2.2. Minería de Datos.</li> <li>2.2.3. Patrones.</li> <li>2.2.4. Evaluación / Interpretación / Visualización.</li> </ul> </li> </ul>
3	Componentes del Entorno de Inteligencia de Negocios.	<ul style="list-style-type: none"> <li>3.1. Orígenes de datos.                             <ul style="list-style-type: none"> <li>3.1.1. Sistemas operacionales.</li> <li>3.1.2. Sistemas Heredados.</li> <li>3.1.3. ERP's, CRM's.</li> <li>3.1.4. Otros.</li> </ul> </li> <li>3.2. Bodegas de datos.                             <ul style="list-style-type: none"> <li>3.2.1. Diseño.</li> <li>3.2.2. Implementación.</li> </ul> </li> <li>3.3. Procesos de ETL.</li> <li>3.4. Procesos de Minería de datos.</li> <li>3.5. Vistas Multidimensionales (hipercubos de datos).</li> <li>3.6. Reportadores.</li> <li>3.7. Alertas, tableros de control (dashboards) e indicadores clave de desempeño (KPI's).</li> <li>3.8. Procesadores de consultas ad-hoc.</li> </ul>
4	Construcción a la solución de la inteligencia de negocios.	<ul style="list-style-type: none"> <li>4.1. Creación del proyecto final integrando las técnicas y herramientas vistas anteriormente tomando en cuenta cada uno de los componentes para la toma de decisiones en la inteligencia de negocios.</li> </ul>

## 7. Actividades de aprendizaje de los temas

1. Introducción a la Inteligencia de Negocios.	
Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"> <li>Identificar los conceptos básicos, herramientas y componentes de la inteligencia de Negocios.</li> </ul> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>Capacidad de análisis y síntesis.</li> <li>Capacidad de organizar y planificar.</li> <li>Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>Solución de problemas.</li> <li>Toma de decisiones.</li> <li>Trabajo en equipo.</li> <li>Capacidad de aplicar los conocimientos.</li> <li>Habilidades de investigación.</li> <li>Capacidad de generar nuevas ideas.</li> <li>Liderazgo.</li> <li>Habilidad para trabajar en forma Autónoma.</li> <li>Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>Investigar la definición de inteligencia de negocios.</li> <li>Discutir el objetivo y las ventajas de la inteligencia de negocios.</li> <li>Analizar las soluciones de la IB desarrollando ejemplos de cada uno de sus componentes.</li> <li>Ejercitar la toma de decisiones.</li> <li>Diseñar el esquema de un almacén de datos (Data Warehouse): que consiga unificar de manera operativa toda la información recogida.</li> <li>Seleccionar y aplicar el método de minería de datos apropiado;</li> <li>Evaluar, interpretar, transformar y representar los patrones extraídos.</li> <li>Tomar decisiones más acertadas para planear los próximos objetivos o corregir alguna desviación a los mismos.</li> </ul>
2. Base de Datos para la Toma de Decisiones.	
Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"> <li>Explotar los esquemas de base de datos multidimensionales utilizando herramientas de visualización, pivoteo y consultas en línea.</li> </ul> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>Capacidad de análisis y síntesis.</li> <li>Capacidad de organizar y planificar.</li> </ul>	<ul style="list-style-type: none"> <li>Identificar ventajas y desventajas para la elección de un sistema de base de datos que den soporte a la toma de decisiones con base a la naturaleza de la información: ¿Qué tipos de datos se usarán: relacional, transaccional, texto, series de tiempo, espacial?</li> <li>Identificar los tipos de escalabilidad de los sistemas de bases de datos: (1) Por Filas (ó tamaño de la base de datos), (2) Por Columnas (ó dimensión)</li> </ul>

<ul style="list-style-type: none"> <li>• Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>• Solución de problemas.</li> <li>• Toma de decisiones.</li> <li>• Trabajo en equipo.</li> <li>• Capacidad de aplicar los conocimientos.</li> <li>• Habilidades de investigación.</li> <li>• Capacidad de generar nuevas ideas.</li> <li>• Liderazgo.</li> <li>• Habilidad para trabajar en forma Autónoma.</li> <li>• Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>• Identificar la configuración del sistema de base de datos: ¿Se utilizará en uno o varios sistemas operativos? ¿Proveerá interfaces basadas en Web y permitirá datos XML como entrada y / o salida? ¿Arquitectura Cliente / Servidor?</li> <li>• Ejemplificar las ventajas y desventajas de ROLAP y MOLAP. Ventajas de MOLAP Mayor rendimiento en el procesamiento de consultas, poco tiempo de cálculo realizado escritura en la base de datos. Desventajas de MOLAP: tamaño limitado en la arquitectura del cubo, no accede a datos que no están en el cubo, no explota el paralelismo. Ventajas de ROLAP.</li> </ul>
---	---

3. Componentes del Entorno de Inteligencia de Negocios.

Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"> <li>• Diseñar e implementar un pequeño datawarehouse o datamart definiendo los metadatos necesarios a utilizar para integrarse a soluciones de inteligencia empresarial haciendo una descripción de los usos y aplicaciones que tiene cada una de ellas.</li> </ul> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>• Capacidad de análisis y síntesis.</li> <li>• Capacidad de organizar y planificar.</li> <li>• Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>• Solución de problemas.</li> <li>• Toma de decisiones.</li> <li>• Trabajo en equipo.</li> <li>• Capacidad de aplicar los conocimientos.</li> <li>• Habilidades de investigación.</li> <li>• Capacidad de generar nuevas ideas.</li> </ul>	<ul style="list-style-type: none"> <li>• Investigar en distintas fuentes de información sobre los componentes que integran una solución de inteligencia de negocios.</li> <li>• Elaborar un diagrama con los componentes de una solución de inteligencia de negocios.</li> <li>• Describir los diferentes orígenes de datos que pueden alimentar a un datawarehouse.</li> <li>• Diseñar e implementar los metadatos que requieren un dataware house.</li> <li>• Buscar y clasificar información sobre tecnologías y herramientas utilizadas para los procesos de ETL's.</li> <li>• Elaborar prácticas donde utilice una herramienta de ETL para mover datos de un sistema operacional a un dataware house.</li> <li>• Elaborar prácticas de diseño de esquemas multidimensionales y probar su funcionalidad con MDX</li> </ul>



<ul style="list-style-type: none"> <li>• Liderazgo.</li> <li>• Habilidad para trabajar en forma. Autónoma.</li> <li>• Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>• Elaborar prácticas donde implemente un algoritmo de minería de datos para obtener algún indicador o patrón.</li> <li>• Elaborar prácticas para elaborar reportes y obtener indicadores clave de desempeño así como tableros de control.</li> </ul>
<p>4. Construcción a la solución de la inteligencia de negocios.</p>	
<p>Competencias</p>	<p>Actividades de aprendizaje</p>
<p>Específica(s):</p> <ul style="list-style-type: none"> <li>• Construir una solución de inteligencia de negocios para un caso práctico.</li> </ul> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>• Capacidad de análisis y síntesis.</li> <li>• Capacidad de organizar y planificar.</li> <li>• Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>• Solución de problemas.</li> <li>• Toma de decisiones.</li> <li>• Trabajo en equipo.</li> <li>• Capacidad de aplicar los conocimientos.</li> <li>• Habilidades de investigación.</li> <li>• Capacidad de generar nuevas ideas.</li> <li>• Liderazgo.</li> <li>• Habilidad para trabajar en forma. Autónoma.</li> <li>• Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>• Crear, desarrollar e integrar un proyecto con las técnicas apropiadas para la toma de decisiones en la inteligencia de negocios.</li> </ul>

## 8. Práctica(s)

- Definir esquemas de bases de datos multidimensionales
- Elaborar consultas a bases de datos multidimensionales
- Diseño e implementación de un datamart o datawarehouse
- Uso de herramientas para la extracción, transformación y carga de datos de una base de datos relacional a una datawarehouse
- Definir y aplicar esquemas de análisis de datos
- Definir y aplicar algoritmos de minería de datos a utilizar
- Desplegar solución
- Elaborar aplicaciones para acceder a la solución desplegada (reportes, consultas mdx, visualización de datos, alertas, conocimiento, etc.).
- Ejemplos de prácticas:
  - Prácticas de filtrado usando alguna herramienta de minería de datos.
  - Prácticas de aprendizaje usando alguna herramienta de minería de datos.
  - Prácticas de meta-aprendizaje usando alguna herramienta de minería de datos.
  - Prácticas de agrupamiento usando alguna herramienta de minería de datos.

Proyecto individual y/o de grupo.

## 9. Proyecto de asignatura

El objetivo del proyecto que planteé el docente que imparta esta asignatura, es demostrar el desarrollo y alcance de la(s) competencia(s) de la asignatura, considerando las siguientes fases:

- **Fundamentación:** marco referencial (teórico, conceptual, contextual, legal) en el cual se fundamenta el proyecto de acuerdo con un diagnóstico realizado, mismo que permite a los estudiantes lograr la comprensión de la realidad o situación objeto de estudio para definir un proceso de intervención o hacer el diseño de un modelo.
- **Planeación:** con base en el diagnóstico en esta fase se realiza el diseño del proyecto por parte de los estudiantes con asesoría del docente; implica planificar un proceso: de intervención empresarial, social o comunitario, el diseño de un modelo, entre otros, según el tipo de proyecto, las actividades a realizar los recursos requeridos y el cronograma de trabajo.
- **Ejecución:** consiste en el desarrollo de la planeación del proyecto realizada por parte de los estudiantes con asesoría del docente, es decir en la intervención (social, empresarial), o construcción del modelo propuesto según el tipo de proyecto, es la fase de mayor duración que implica el desempeño de las competencias genéricas y específicas a desarrollar.
- **Evaluación:** es la fase final que aplica un juicio de valor en el contexto laboral-profesión, social e investigativo, ésta se debe realizar a través del reconocimiento de logros y aspectos a mejorar se estará promoviendo el concepto de “evaluación para la mejora continua”, la metacognición, el desarrollo del pensamiento crítico y reflexivo en los estudiantes.

## 10. Evaluación por competencias

Para evaluar las actividades de aprendizaje se recomienda solicitar: mapas conceptuales, reportes de prácticas, estudios de casos, exposiciones en clase, ensayos, problemarios, reportes de visitas, portafolio de evidencias y cuestionarios, cuadro sinóptico.

Para verificar el nivel del logro de las competencias del estudiante se recomienda utilizar: listas de cotejo, listas de verificación, matrices de valoración, guías de observación, coevaluación y autoevaluación.

## 11. Fuentes de información

Impresas:

1. Margaret h. Dunham. Data mining: introductory and advanced topics. Prentice hall.
2. J. Han and m. Kamber. Data mining: concepts and techniques. Morgan kaufmann Publishers.
3. D. Hand, h. Mannila and p. Smyth. Principles of data mining. Mit press.
4. Ian h. Witten, eibe frank. Data mining: practical machine learning tools and Techniques with java implementations. Morgan kaufmann publishers.
5. Mehmed kantardzic. Data mining: concepts, models, methods, and algorithms wiley-IEEE press.
6. Tan, steinbach, kumar. Introduction to data mining. Addison-wesley.
7. Usama m. Fayyad, gregory piatetsky-shapiro, padhraic smyth. Advances in Knowledge discovery and data mining. Aai/mit press.
8. Ian h. Witten, eibe frank. Data mining: practical machine learning tools and Techniques. Morgan kaufmann, 2nd edition.
9. Dorian pyle. Data preparation for data mining. Morgan kaufmann.
10. daniel t. Larose. Discovering knowledge in data : an introduction to data mining. Wiley-interscience.
11. mastering data warehouse design - relational and dimensional techniques. Ed Wiley. 2003.
12. data analysis -the data warehouse toolkit - second edition. Ed wiley.
13. building the data warehouse – third edition. Ed wiley.
14. the data warehouse etl toolkit. Ed wiley 2005.
15. the data warehouse lifecycle toolkit. Ed wiley. 1998.