

1.- DATOS DE LA ASIGNATURA

Nombre de la asignatura: Minería de datos
Carrera: Ingeniería en Sistemas Computacionales
Clave de la asignatura: BDE - 0803
Horas teoría-horas práctica-créditos: 2-2-6

2.- HISTORIA DEL PROGRAMA

Lugar y fecha de elaboración o revisión	Participantes	Observaciones (cambios y justificación)
Instituto Tecnológico de Acapulco del 6 al 10 de agosto de 2007	Academia de Sistemas y Computación (Ing. Ricardo Flores Oliveros, Ing. Eleazar Pacheco Salazar, Ing. Juan José Bedolla Solano).	Definición de los programas de estudio de la carrera de Ingeniería en Sistemas Computacionales.

3.- UBICACIÓN DE LA ASIGNATURA

a) Relación con otras asignaturas del plan de estudio

Anteriores	
Asignaturas	Temas
Fundamentos de base de datos Programación I, II, III y IV	Álgebra relacional Algoritmos

Posteriores	
Asignaturas	Temas
Procesamiento analítico en línea Data warehouse	Consultas de lógica difusa Modelos multidimensionales

b). Aportación de la asignatura al perfil del egresado

Modela y diseña los modelos multidimensionales de datos.

4.- OBJETIVO(S) GENERAL(ES) DEL CURSO

El estudiante comprenderá los modelos multidimensionales de datos, así como las técnicas de selección, procesamiento y extracción de la información.

5.- TEMARIO

Unidad	Temas	Subtemas
1	Minería de datos.	1.1. Qué es minería de datos. 1.2. Tipos de datos. 1.3. Tipos de modelos de datos. 1.4. Proceso de descubrimiento de conocimiento. 1.5. Sistemas para el soporte de decisiones. 1.6. Sistemas operacionales vs Sistemas para la toma de decisiones. 1.7. Herramientas para la minería de datos.
2	Proceso de extracción.	2.1. Introducción. 2.2. Etapas del proceso de extracción del conocimiento. 2.3. Etapa de recopilación e integración. 2.4. Etapa de selección, limpieza y transformación. 2.5. Etapa de minería de datos. 2.6. Etapa de evaluación e interpretación. 2.7. Etapa de difusión.

3	Sistemas de bases de datos multidimensionales.	3.1 Introducción. 3.2. Almacenes de datos. 3.3. Estructura de los almacenes de datos. 3.4. Arquitectura de los almacenes de datos. 3.5. Carga y mantenimiento de los datos.
4	Selección de los datos.	4.1. Preparación de las vistas. 4.2. Qué son las consultas de lógica difusa. 4.3. Selección de lenguajes. 4.4. Primitivas. 4.5. Interfaces de minería de datos.
5	Extracción de datos.	5.1. Aprendizaje inductivo. 5.2. Definición de Patrones. 5.3. Aplicación de consultas con lógica difusa. 5.4. Cubos OLAP.

6.- APRENDIZAJES REQUERIDOS

- Tener conocimientos previos de diferentes modelos de bases de datos de acuerdo a su estructura.
- Tener nociones de los modelos multidimensionales.

7.- SUGERENCIAS DIDÁCTICAS

- **Propiciar la investigación previa a la clase de los conceptos de la asignatura para su análisis y discusión en grupo.**
- **Propiciar el estudio individual y obtener conclusiones grupales.**
- **En equipo, realizar exposiciones.**
- **Presentar en equipo proyectos finales.**
- **Exponer en equipo los resultados de los proyectos finales.**

8.- SUGERENCIAS DE EVALUACIÓN

- **Participación en forma individual y grupal.**
- **Exámenes departamentales.**
- **Desarrollo de un proyecto final aplicando los conocimientos adquiridos en las unidades de aprendizaje.**
- **Evaluación de la presentación del proyecto final (Informe, presentación, defensa y congruencia del proyecto).**

9.- UNIDADES DE APRENDIZAJE

UNIDAD 1.- Minería de datos.

Objetivo Educativo	Actividades de Aprendizaje	Fuentes de Información
El estudiante identificará conceptos básicos de la minería de datos.	1.1 Buscar información sobre minería de datos. 1.2 Discutir los diferentes modelos de datos. 1.3 Comparar los sistemas operacionales con los sistemas para la toma de decisiones. 1.4 Identificar las herramientas para la minería de datos.	1, 2, 5 11

UNIDAD 2.- Proceso de extracción.

Objetivo Educativo	Actividades de Aprendizaje	Fuentes de Información
Conocerá las diferentes etapas del proceso de extracción y analizará los diferentes algoritmos y técnicas de	2.1 Buscar y analizar las etapas del proceso de extracción del conocimiento. 2.2 Discutir las etapas de recopilación e integración, de selección, limpieza y transformación, así como las de	3, 4, 6 12

selección, procesamiento y extracción de datos.	evaluación e interpretación y difusión.	
-------------------------------------------------	------------------------------------------------	--

UNIDAD 3.- Sistemas de bases de datos multidimensionales.

Objetivo Educativo	Actividades de Aprendizaje	Fuentes de Información
Conocerá la estructura, arquitectura, carga y mantenimiento de los almacenes de datos.	3.1 Investigar conceptos de almacenes de datos. 3.2 Analizar la estructura y arquitectura de los almacenes de datos. 3.3 Discutir como se lleva a cabo la carga y mantenimiento de los datos.	7, 9 10

UNIDAD 4.- Selección de los datos.

Objetivo Educativo	Actividades de Aprendizaje	Fuentes de Información
Entenderá la lógica difusa y generación de vistas de selección.	4.1 Investigar que son las consultas de lógica difusa. 4.2 Analizar la selección de lenguajes. 4.3 Conocer las interfaces de minería de datos.	7, 8 13

UNIDAD 5.- Extracción de datos.

Objetivo Educativo	Actividades de Aprendizaje	Fuentes de Información
Realizará un estudio sobre lógica difusa, modelos multidimensionales e inteligencia artificial.	5.1 Entender que es el aprendizaje inductivo. 5.2 Realizar ejercicios de aplicación de consultas con lógica difusa. 5.3 Investigar acerca de los cubos OLAP.	7, 12 13, 14

10.- FUENTES DE INFORMACIÓN

1. Ullman, J.D., *Principles of dDatabases and Knowledge Base Systems*. 1992.
2. Pyle D. Data Preparation For Fata Mining
3. Berry J.A., Linoff S. Mastering Data Mining
4. Doug Wielenga, Bob Lucas & Jim Georges, Enterprise Miner: Applying Data Mining Techniques Course Notes, SAS Institute Inc. Cary, Nc, USA. 1999
5. Ramez A. Elmasri, S.B.N., *Fundamentos de Sistemas de Bases de Datos*. Tercera Edición ed. Vol. 1. 2002, Madrid.
6. A.F. Cardenas, D.M., *Research Fandations and Semantic Database Systems*. 1992.
7. Thomsen, E., *OLAP Solutions - Building Multidimensional Information Systems*. 1997.
8. Ralph Kimball, L.R., Margy Ross, Warren Thornthwaite, *The Data Warehouse Lifecycle Toolkit*. First Edition ed. 1998, States United of America.
9. Matthias Jarke, M.L., Yannis Vassiliou, Panos Vassiliadis, *fundamentals of Data Warehouses*. Vol. 1. Sept. 2002, Berling. 213.
10. Inmon, W.H., *Building the Data Warehouse*. Second ed. Vol. 1. 1976, United States of America. 401.
- 11.. R. and Srikant, R., *Fast Algorithms for Mining Association Rules in Large Databases*. Second edition ed. 1994.
12. Albert, J., *Data integration in the RODIN multidatabase system*. 1996, USA.
13. Group, A., *Data Warehouse Query Tools:Evolving to relational OLAP*. 1995.
14. Vidette Poe, P.K., Stephen Brobst., *Building a Data Warehouse for Decision Support*. 1998.

11.- PRÁCTICAS SUGERIDAS

MATERIA: MINERÍA DE DATOS	CLAVE:	EQUIPO NO:
PROFESOR: RICARDO FLORES OLIVEROS, ELEAZAR PACHECO SALAZAR JUAN JOSÉ BEDOLLA SOLANO	FECHA:26-JUN-2008	
PRÁCTICA #1: Algoritmos de árboles de decisión		
OBJETIVO (S): EL alumno, reforzará los conocimientos sobre regresión y los atributos de entrada utilizando los valores de entrada o de estados de las columnas de datos.		

MATERIAL:

- SQL Server 2005 Analysis Services (SSAS) para el modelado de predicción de atributos discretos y continuos.

EQUIPO:

- Computadora personal.
- Un sistema operativo Windows de 32 bits, XP
- Se recomienda fuertemente el uso de un sistema operativo Windows basado en NT (NT, 2000, XP, 2003) puesto que éstos permiten ejecutar el servidor MySQL como un servicio.

HERRAMIENTA:

- Microsoft SQL Server 2005 Analysis Services (SSAS)

DIAGRAMAS:

El alumno generará las gráficas necesarias como resultados

ACTIVIDADES:

1. Generar el modelo de datos.
2. Seleccionar los datos
3. Alimentar el modelo con los datos
4. Hacer uso de la herramienta Microsoft SQL Server 2005 Analysis Services (SSAS)

BIBLIOGRAFÍA:

- Thomsen, E., *OLAP Solutions - Building Multidimensional Information Systems*. 1997.
- <http://msdn.microsoft.com/es-es/library/ms167488.aspx>

CONTENIDO DEL REPORTE:

1. visualización de gráficas y resultados en pantalla

MATERIA: MINERÍA DE DATOS	CLAVE:	EQUIPO NO:
PROFESOR: RICARDO FLORES OLIVEROS, ELEAZAR PACHECO SALAZAR JUAN JOSÉ BEDOLLA SOLANO	FECHA: 26-JUN-2008	
PRÁCTICA #2: Algoritmo de clústeres.		
OBJETIVO (S): El alumno, reforzará el conocimiento sobre clusterización		

MATERIAL:

- SQL Server 2005 Analysis Services (SSAS)

EQUIPO:

- Computadora personal.
- Un sistema operativo Windows de 32 bits, XP
- Se recomienda fuertemente el uso de un sistema operativo Windows basado en NT (NT, 2000, XP, 2003) puesto que éstos permiten ejecutar el servidor MySQL como un servicio.

HERRAMIENTA: Microsoft SQL Server 2005 Analysis Services (SSAS)

DIAGRAMAS: El alumno generará las gráficas necesarias como resultados

ACTIVIDADES:

5. Generar el modelo de datos.
6. Seleccionar los datos
7. Alimentar el modelo con los datos
8. Hacer uso de la herramienta Microsoft SQL Server 2005 Analysis Services (SSAS)

BIBLIOGRAFÍA:

1. Pyle D. Data Preparation For Fata Mining
2. Berry J.A., Linoff S. Mastering Data Mining
3. Doug Wielenga, Bob Lucas & Jim Georges, Enterprise Miner: Applying Data Mining Techniques Course Notes, SAS Institute Inc. Cary, Nc, USA. 1999
4. Ramez A. Elmasri, S.B.N., *Fundamentos de Sistemas de Bases de Datos*. Tercera Edición ed. Vol. 1. 2002, Madrid.

CONTENIDO DEL REPORTE:

1. Revisión de los resultados en la computadora

MATERIA: MINERÍA DE DATOS	CLAVE:	EQUIPO NO:
---------------------------	--------	------------

PROFESOR: RICARDO FLORES OLIVEROS, ELEAZAR PACHECO SALAZAR JUAN JOSÉ BEDOLLA SOLANO	FECHA:26-JUN-2008
-------------------------------------------------------------------------------------------	-------------------

PRÁCTICA #3: Predicción de los datos

OBJETIVO (S): El alumno, será capaz de proporcionar los datos suficientes que sirvan para la predicción considerando las columnas de entra y de predicción de manera independiente.

MATERIAL:

- Microsoft SQL Server 2005 Analysis Services (SSAS)

EQUIPO:

- Computadora personal.
- Un sistema operativo Windows de 32 bits, XP
- Se recomienda fuertemente el uso de un sistema operativo Windows basado en NT (NT, 2000, XP, 2003) puesto que éstos permiten ejecutar el servidor MySQL como un servicio.

HERRAMIENTA:

: Microsoft SQL Server 2005 Analysis Services (SSAS)

DIAGRAMAS: El alumno generará las gráficas necesarias como resultados

ACTIVIDADES:

9. Generar el modelo de datos.
10. Seleccionar los datos
11. Alimentar el modelo con los datos
12. Hacer uso de la herramienta Microsoft SQL Server 2005 Analysis Services (SSAS)

BIBLIOGRAFÍA:

5. Pyle D. Data Preparation For Fata Mining
6. Berry J.A., Linoff S. Mastering Data Mining
7. Doug Wielenga, Bob Lucas & Jim Georges, Enterprise Miner: Applying Data Mining Techniques Course Notes, SAS Institute Inc. Cary, Nc, USA. 1999
8. Ramez A. Elmasri, S.B.N., *Fundamentos de Sistemas de Bases de Datos*. Tercera Edición ed. Vol. 1. 2002, Madrid.

CONTENIDO DEL REPORTE:

2. Revisión de los resultados en la computadora

MATERIA: MINERÍA DE DATOS	CLAVE:	EQUIPO NO:
PROFESOR: RICARDO FLORES OLIVEROS, ELEAZAR PACHECO SALAZAR JUAN JOSÉ BEDOLLA SOLANO	FECHA:26-JUN-2008	
PRÁCTICA #4: Predicción de asociación		
OBJETIVO (S): El alumno, será capaz de proporcionar los datos suficientes que sirvan para la asociación considerando las columnas de entra y de predicción de manera independiente.		

MATERIAL:

- Microsoft SQL Server 2005 Analysis Services (SSAS)

EQUIPO:

- Computadora personal.
- Un sistema operativo Windows de 32 bits, XP
- Se recomienda fuertemente el uso de un sistema operativo Windows basado en NT (NT, 2000, XP, 2003) puesto que éstos permiten ejecutar el servidor MySQL como un servicio.

HERRAMIENTA:

: Microsoft SQL Server 2005 Analysis Services (SSAS)

DIAGRAMAS: El alumno generará las gráficas necesarias como resultados

ACTIVIDADES:

13. Generar el modelo de datos.
14. Seleccionar los datos
15. Alimentar el modelo con los datos
16. Hacer uso de la herramienta Microsoft SQL Server 2005 Analysis Services (SSAS)

BIBLIOGRAFÍA:

9. Pyle D. Data Preparation For Fata Mining
10. Berry J.A., Linoff S. Mastering Data Mining
11. Doug Wielenga, Bob Lucas & Jim Georges, Enterprise Miner: Applying Data Mining Techniques Course Notes, SAS Institute Inc. Cary, Nc, USA. 1999
12. Ramez A. Elmasri, S.B.N., *Fundamentos de Sistemas de Bases de Datos*. Tercera Edición ed. Vol. 1. 2002, Madrid.

CONTENIDO DEL REPORTE:

3. Revisión de los resultados en la computadora

MATERIA: MINERÍA DE DATOS	CLAVE:	EQUIPO NO:
PROFESOR: RICARDO FLORES OLIVEROS, ELEAZAR PACHECO SALAZAR JUAN JOSÉ BEDOLLA SOLANO	FECHA:26-JUN-2008	
PRÁCTICA #5: Selección de características en minería de datos		
OBJETIVO (S): El alumno, será capaz de proporcionar los datos suficientes que sirvan para la selección de características en minería de datos .		

MATERIAL:

- Microsoft SQL Server 2005 Analysis Services (SSAS)

EQUIPO:

- Computadora personal.
- Un sistema operativo Windows de 32 bits, XP
- Se recomienda fuertemente el uso de un sistema operativo Windows basado en NT (NT, 2000, XP, 2003) puesto que éstos permiten ejecutar el servidor MySQL como un servicio.

HERRAMIENTA:

: Microsoft SQL Server 2005 Analysis Services (SSAS)

DIAGRAMAS: El alumno generará las gráficas necesarias como resultados

ACTIVIDADES:

17. Generar el modelo de datos.
18. Seleccionar los datos
19. Alimentar el modelo con los datos
20. Hacer uso de la herramienta Microsoft SQL Server 2005 Analysis Services (SSAS)

BIBLIOGRAFÍA:

13. Pyle D. Data Preparation For Data Mining
14. Berry J.A., Linoff S. Mastering Data Mining
15. Doug Wielenga, Bob Lucas & Jim Georges, Enterprise Miner: Applying Data Mining Techniques Course Notes, SAS Institute Inc. Cary, Nc, USA. 1999
16. Ramez A. Elmasri, S.B.N., *Fundamentos de Sistemas de Bases de Datos*. Tercera Edición ed. Vol. 1. 2002, Madrid.

CONTENIDO DEL REPORTE:

4. Revisión de los resultados en la computadora