

UNIVERSIDAD NACIONAL DE ASUNCIÓN
FACULTAD POLITÉCNICA
PROGRAMA DE ESTUDIO
INGENIERIA EN INFORMÁTICA

Consejo Directivo - Resolución 13/05/18-00 Acta 876/11/03/2013 Anexo 07

I. - IDENTIFICACIÓN

1. Materia : Electiva 5 – Data Warehouse | Diseño Multidimensional
2. Semestre : Noveno
3. Horas semanales : 7 horas
 - 3.1. Clases teóricas: 7 horas
 - 3.2. Clases prácticas: 0 horas
4. Total real de horas disponibles: 112 horas
 - 4.1. Clases teóricas: 98 horas
 - 4.2. Clases prácticas: 14 horas

II. - OBJETIVO

El principal objetivo de esta asignatura es capacitar al estudiante para reconocer el potencial y la problemática del análisis de sistemas de información para la toma de decisiones, conocer el modelo multidimensional de datos de los almacenes de datos y los operadores de refinamiento asociados. Del mismo modo, se busca que el estudiante comprenda las fases del descubrimiento de conocimiento de Bases de Datos y la importancia de las mismas en el éxito del proceso. Por último, el estudiante deberá reconocer los beneficios del uso sistemático de técnicas de extracción de conocimiento para la obtención de modelos y patrones predictivos o descriptivos.

III. - PRE-REQUISITO

El alumno debe haber aprobado el 70% de los créditos de la carrera y 300 horas de pasantía para cursar la Electiva 5 de acuerdo al Plan 2008 de la Carrera de Ingeniería en Informática.

IV. - CONTENIDO

Tema 1: Introducción a Almacenes de Datos y Minería de Datos

1. Definiciones de almacenes de datos. Motivación y características. Arquitectura.
2. Sistemas ROLAP (*Relational Online Analytical Processing*) y MOLAP (*Multidimensional OnLine Analytical Processing*). Carga y Mantenimiento de un Almacén de Datos. Diseño de un almacén de datos.
3. Introducción a la Minería de Datos. Motivación. Problemas tipo y aplicaciones.
4. Técnicas de Minería de Datos.
5. Líneas de investigación abiertas.

Tema 2: Planeamiento de Proyectos de Almacenes de Datos

1. Marco dimensional del ciclo de vida del negocio.
2. Preparación y evaluación de riesgos.
3. Priorización de alcance y justificaciones.
4. Roles y responsabilidades del equipo de desarrollo.
5. Desarrollo y mantenimiento del plan de proyecto.
6. Definición de los requisitos del negocio.
7. Técnicas para la recolección de requisitos.

Tema 3: Modelado Dimensional

1. El rol del modelado dimensional.
2. Características de la tabla del hecho y de la dimensional.
3. Cuatro pasos para diseñar modelos dimensionales.
4. Des-normalización de jerarquías dimensionales. Llaves sustitutas.
5. Tablas del hecho sin hechos.
6. Variaciones del copo de nieve.
7. Dimensiones degeneradas.
8. Implicancias de cadena de valores.
9. Arquitectura de bus y la matriz de integración de modelos dimensionales.
10. Dimensiones confortantes.
11. El rol de las dimensiones.
12. Hechos semi aditivos y no aditivos.
13. Productos heterogéneos.
14. Consideraciones para el diseño de numerosas dimensiones.

Tema 4: Arquitectura de los Almacenes de Datos

1. Definiciones y descripciones.
2. Arquitectura del Almacenamiento de Datos.
3. Modelos, niveles de detalle.
4. Componentes y servicios comunes.



5. Sitio Trasero.
6. Consideraciones sobre el área de presentación.
7. Sitio Delantero.
8. Tipos de usuarios del almacén de los datos.
9. Reportes estándar.
10. Minera de Datos.
11. Depósito de los meta datos.
12. Contenido. Estándares – XML (*Extensible Markup Language*).
13. El proceso de desarrollo de la arquitectura en ocho pasos

Tema 5: Opciones de Configuración Física

1. Infraestructura.
2. Almacén de los datos vs. Data Marts.
3. Data Marts aislados.
4. Almacén de Datos empresariales.
5. Almacén de Datos confortantes.
6. Tipos de almacenes de datos operacionales (ODS - *Operational Data Store*).
7. Arquitecturas de presentación – OLAP (*OnLine Analytical Processing*), ROLAP (*Relational Online Analytical Processing*), MOLAP (*Multidimensional Online Analytical Processing*), HOLAP (*Hybrid Online Analytical Process*).

Tema 6: Selección de Productos

1. Proceso de selección basado en la arquitectura
2. Matriz de evaluación de productos

Tema 7: Diseño Físico de la Base de Datos

1. Estándares y convenciones de nombramiento.
2. Desarrollo del modelo físico.
3. Navegación y estrategia de la agregación.
4. Tipos y pautas de índices.
5. Gerencia de datos.

Tema 8: Procesos y Técnicas de Preparación de Datos

1. Diseño de la aplicación para preparación de los datos.
2. Preparación de las tablas de dimensionales.
3. Preparación de las tablas hechos.
4. Validación de procesos.
5. Implementación de las operaciones del almacén de datos.
6. Técnicas básicas de extracción y transformación de datos.
7. Creación y mantenimiento de claves sustitutas.
8. Aceleración del ciclo de carga de dato.
9. Integración de datos desde múltiples fuentes.
10. Calidad y limpieza de los datos

Tema 9: Despliegue y Perspectiva de Crecimiento

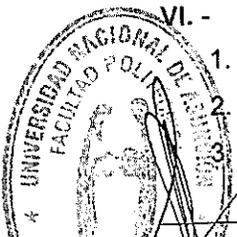
1. Rol de las aplicaciones para usuarios finales.
2. Diseño de plantilla de aplicaciones y navegación.
3. Proceso de entrega de la aplicación.
4. Soporte y entrenamiento.
5. Documentación.
6. Planes para el crecimiento.

V. - ESTRATEGIAS METODOLÓGICAS.

1. Exposición oral.
2. Laboratorio.
3. Formación de grupos y trabajo en equipo para resolver problemas durante las clases.
4. Trabajos prácticos.

VI. - MEDIOS AUXILIARES

1. Pizarra.
2. Proyector Multimedia.
3. Sala de máquinas para clases prácticas.



4. Bibliografía de apoyo.

VII. - EVALUACIÓN

La evaluación se realizará de acuerdo a las reglamentaciones de la Facultad y la Carrera.

VIII. - BIBLIOGRAFÍA

Bibliografía Básica

1. The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing, and Deploying Data Warehouses, Ralph Kimball. 2008.

Bibliografía Complementaria

1. The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning by Ralph Kimball and Joe Caserta, 2004.

