Lectures for the course: Data Warehousing and Data Mining (IT 60107)

Week 1

Lecture 1 – 24/07/2006

• Introduction to the course
• Expectations
• Evaluation Guideline
• Term Paper and Term Project Guideline

Lecture 2(A+B) – 26/07/2006

• What is a Data warehouse
• Characteristics of OLTP applications
• Why a Data Warehouse is required

Lecture 3 – 27/07/2006

• What is a Data warehouse – contd.
• Characteristics of Data Warehouse applications
• De-normalized Database Design

Week 2

Lecture 4 – 31/07/2006

• Data Warehouse Architecture
• Operational Data Sources, Data Staging Area, Data Warehouse DB, OLAP Server and Client Tools

Lecture 5(A+B) – 02/08/2006

• OLAP Operations
• Slicing, Dicing, Roll-up, Drill-down

Lecture 6 – 03/08/2006

• Relational Schema Design for supporting OLAP operations

Week 3
Lecture 7 –07/08/2006

- Supporting OLAP operations using Relational Schema
- Multi dimensional array representation of data warehouse data
- ROLAP, MOLAP and HOLAP
- Lattice of cuboids
- Possible number of summarized tables

Lecture 8(A+B) –09/08/2006

- Dimension hierarchies
- Number of cuboids in a lattice with dimension hierarchies
- Roll-up and Rill down by dimension hierarchies
- Materialized Views

Lecture 9 –10/08/2006

- Introduction to the view materialization problem

Week 4

Lecture 7 –14/08/2006

- A greedy algorithm for selecting materialized views

Lecture 8(A+B) –16/08/2006

- Computation of multidimensional arrays from chunks

Lecture 9 –17/08/2006

- Star Schema
- Retail Sales Data Warehouse Schema Design
- Size estimation of Fact table and Dimension tables
- Size estimation of fact tables

Week 5

Lecture 10 –21/08/2006

- More insight into Retail Sales Schema Design
- Product Dimension
- Date Dimension
- Store Dimension
- Normalized fact table
• Degenerate Dimension
• De-normalized dimension tables

**Lecture 11 (A+B) – 23/08/2006**

• Snowflake Schema
• Fact Constellation schema
• Transaction No.
• Extensibility of Star Schema
• Effect of addition of new attributes in the dimensional tables

**Lecture 12 – 24/08/2006**

• Extensibility of Star Schema contd.
• Effect of change in granularity
• Effect of adding new facts
• Time dimension

**Week 6**

**Lecture 13 – 28/08/2006**

• Fact less Fact Table

**Lecture 14 (A+B) – 30/08/2006**

• Class Test 1 held here

**Lecture 15 – 31/08/2006**

• Class test script shown

**Week 7**

**Lecture 16 – 04/09/2006**

• Inventory business process
• Periodic Snapshot Schema

**Lecture 17 (A+B) – 06/09/2006**

• Inventory Transaction Schema
• Slowly changing dimension
Lecture 18 – 07/09/2006

- Slowly changing dimension – contd.
- Rapidly changing dimension

Week 8


- Rapidly changing dimension – contd.
- Selective normalization

Lecture 20 (A+B) –13/09/2006

- Summary

Lecture 21 – 14/09/2006

- No Class

Week 9

Mid sem was held here

Week 10


- Introduction to Data Mining
- Association Rule
- Itemset
- Support

Lecture 23 (A+B) –27/09/2006

- A priori Algorithm
- Confidence

Week 11

Autumn Break

Week 12
Lecture 24 – 09/10/2006

- Partitioning Approach

Lecture 25 (A+B) –11/10/2006

- Dynamic Itemset counting approach

Lecture 26 – 12/10/2006

- FP Tree Construction and Mining

Week 13

Lecture 27 – 16/10/2006

- FP Tree Construction and Mining – Running Example

Lecture 28 (A+B) –18/10/2006

- Sequential Pattern Mining

Lecture 29 – 19/10/2006

- Sequential Pattern Mining (contd.)

Week 14

Lecture 30 – 23/10/2006

- Sequential Pattern Mining (contd.)

Lecture 31 (A+B) –25/10/2006

- Clustering - Introduction

Lecture 32 – 26/10/2006

- K-Means Clustering
  - Introduction to k-medoid clustering

Week 15

Lecture 33 – 30/10/2006

- PAM
• CLARA

Lecture 34(A+B) – 01/11/2006

• Clustering – CLARANS
• Introduction to Hierarchical Clustering

Lecture 35 – 02/11/2006

• No Class

Week 16

Lecture 36 – 06/11/2006

• Agglomerative Hierarchical Clustering

Lecture 37(A+B) – 08/11/2006

• BIRCH

Lecture 38 – 09/11/2006

• Classification – Introduction
• Decision Tree

Week 17


• Building Decision Trees


• MLP Based Classification

Lecture 41 – 16/11/2006

• Revision and Summary