INT323:DATABASE ESSENTIALS TOWARD INFORMATICA

L:2 T:0 P:2 Credits:3

Course Outcomes: Through this course students should be able to

CO1 :: infer the usage and application of databases in real world

CO2 :: discover the technology and tools to implement a successful data warehouse

CO3 :: outline the methods and tools for extracting, transforming and loading data

CO4 :: experiment with all common data preparations and transformations for data mining

CO5 :: evaluate model quality with respect to several different performance criteria

CO6 :: make use of data management for business perspectives

Unit I

Digital Data : Introduction to data, Types of data: Structured data, Semi structured Data, Unstructured data, Limitations of databases in the real world, Introduction to OLTP and OLAP

Business Intelligence(BI) : Introduction to BI, Evolution of BI and role of EIS, MIS and Digital Dashboards, Need for BI at all levels, BI for past, present and future

Unit II

Basics of Data Integration : Introduction to data integration: Approaches, need and advantages, technologies used, Introduction to Data Warehouse, data mart, data quality

Unit III

Data Warehouse : Overview, Features of data warehouse, applications, types of data warehouse, functions of data warehouse, tools and utilities, Architecture, OLAP and its types, schemas, meta-data, Data-marting, Security and Backup

Unit IV

Multi-dimensional data : ETL Overview, extracting data, transformations, loading data, simple ETL processing, ETL tools, data cube

Unit V

ETL processing with SQL Server Integration Services : Introduction to SSIS ecosystem, Getting started: User Interface, Creating SSIS projects, data flow and task flow, performing different transformations on ETL jobs

ETL processing with Rapid Miner : Introduction to Rapid Miner ecosystem, Getting started: User interface, creating and managing repositories, operations and processes, Storing data, Processes and result sets, EDL(Exploratory Data Analysis)- Loading Data, Summary statistics, Visualizing data and basic charting, Data Preparation: Basic Data ETL, data types and performing transformation of values, handling missing values, filtering, normalization and standardization

Unit VI

Cloud Databases : Introduction to cloud databases, Cloud databases Vs Traditional database, Benefits and Features of cloud databases, Considerations for Cloud databases, Cloud Providers

Master Data Management(MDM) : Introduction to Master data, Need of managing master data, Introduction to master data management, Key Functions and Benefits of MDM

List of Practicals / Experiments:

List of Practicals/Experiments

- 1. Practical on Data Integration and its approaches
- 2. Practical on OLAP Operations
- 3. Practical on Extract Transform Load (ETL) process
- 4. Setting Up Rapid Miner
- 5. Exploratory Data Analysis
- 6. Practical on Data Preparation, handling missing values
- 7. Cloud Databases

References:

1. DATA WAREHOUSING, DATA MINING, & OLAP by ALEX BERSON, STEPHEN SMITH, MCGRAW HILL EDUCATION

2. DATA MINING: CONCEPTS AND TECHNIQUES by JIAWEI HAN, MICHELINE KAMBER, JIAN PEI, MORGAN KAUFMANN

3. FUNDAMENTALS OF BUSINESS ANALYTICS by RN PRASAD SEEMA ACHARYA, Wiley India