



Excelssior Education Society's
K.C. College of Engineering and Management Studies and Research
Mith Bunder Road, Near Hume Pipe, Kopri, Thane (E)-400603

Name of the event: Guest Lecture on Data Warehousing Mining and Distributed File System

Name of Speaker: Mrs. Renuka Parab

Date: 03/03/2016

Guest lecture on Data Warehousing and Mining & Distributed File System by Mrs. Renuka Parab (Mikado Solutions, Thane) 2015-16
Date: 3rd March 2016



K.C. College of Engineering & Management

DATA WAREHOUSING MINING AND DISTRIBUTED FILE SYSTEM

Topic : Data Warehousing Mining and Distributed File Systems

Speaker Name : Mrs. Renuka Prasad Parab

Target Audience : B.E computer Engineering

Seminar Date : 3/03/2016

Venue : Seminar Hall.

Objectives:To apply the knowledge of Data Warehousing and Distributed file system in IT industry.

Content:

1. What is Big Data: Data which is beyond the storage capacity and stored externally in data center. It is defined by volume , velocity and variety of data.
- 2 .Examples of Big Data in real world.:data generated by facebook, newyork stock exchange,flight details, electricity consumed data.
3. Challenges in big data.:analysis, capturing sharing, storage , transfer, visualization and querying
- 4 What is Data Warehouse: Collection of data from different heterogenous sources like ppt log files , rdbms ,CSV word.It can be structured and semistructured.
5. Data Warehouse Architecture: The entire block diagram was explained which included ETL(Extract Transform Loading). Extraction can be full ie only data is extracted from database or it can be partial ie only updated and current data is extracted from database. In transformation duplication of data is eliminated.
6. Hadoop File system: To store big data in data warehouse we have solution which is Hadoop File System.It is a ready made framework to store big data
7. What is distributed file system: Data is stored at different locations but appears data arriving from a single system to all.

8. Hadoop distributed file system: Detailed explanation on Name node, secondary name node, job tracker, data node and task tracker was done

9. Map reduce :It is a programming model for mapping data and reducing data which is used for processing of data

Topic most liked by Student: Hadoop File system

Topic: Data Warehousing Mining and Distributed File Systems

Objective:

To study the concepts of Data Warehousing Mining and Distributed File Systems

Speaker Name: Mrs. Pooja Prasad Nayak

Experience of speaker: 10 Yrs Teaching & 2 Yrs Industry Experience

Duration: 1 Hr

Venue: Seminar Hall 308 classroom

Target Student: B.E (Computer Engineering)

Prepared by: Kirti Dinesh Kharatrol