Jyoti Nivas College Autonomous
Post Graduate Centre

Database Management System

E-Journal - Department of MCA

Issue: 6 **February 2018**

INDEX

DECISION TREE IN DATA MINING TECHNIQUE

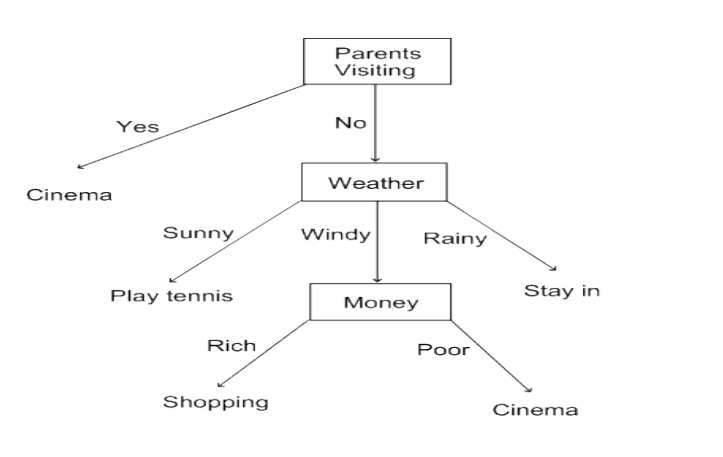
ARSHIYA ANJUM (17MCA07)

K.NISHA (17MCA14)

INTRODUCTION:

The decision tree is one of the most commonly used data mining techniques because its model is easy to understand for users. In decision tree structure that includes a root node, branches, and leaf nodes. The root of the decision tree is a simple question or condition that has multiple answers. Each internal node denotes a test on an attribute, each branch denotes the outcome of a test, and each leaf node holds a class label. Decision trees provides results that can be easily understood by the user.

Example:



You are making your weekend plans and find out that your parents might come to town. You'd like to have plans in place, but there are a few unknown factors that will determine what you can, and can't, do. Time for a decision tree.

First, you draw your decision box. This is the box that includes the event that starts your decision tree. In this case it is your parents coming to town. Out of that box, you have a branch for each possible outcome. In our example, it's easy: yes or no - either your parents come or they don't.

Your parents love the movies, so if they come to town, you'll go to the cinema. Since the goal of the decision tree is to decide your weekend plans, you have an answer. But, what about if your parents don't come to town? We can go back up to the 'no branch' from the decision box and finish our decision tree.

If your parents don't come to town, you need to decide what you are going to do. As you think of options, you realize the weather is an important factor. Weather becomes your next box. Since its spring time, you know it will either be rainy, sunny, or windy. Those three possibilities become your branches.

ADVANTAGES:

- It does not require any domain knowledge.
- It is easy to comprehend.
- The learning and classification steps of a decision tree are simple and fast

APPLICATIONS:

- Decision tree technique is mostly used by statisticians to find out which database is more related to the problem of the business.
- Decision tree technique can be used for Prediction and Data pre-processing.

DISADVANTAGES:

- If the segment contains only one record
- All the records contain identical features
- The growth is not enough to make any further spilt

REFERENCES:

http://www.zentut.com/data-mining/data-mining-techniques/

https://www.tutorialspoint.com/data_mining/dm_dti.htm

https://www.tutorialspoint.com/data_mining/dm_dti.htm

DIGITIZATION OF LIBRARY RESOURCES
DIGITAL LIBRARIES

VARSHA SAHAY (17MCA28)

APURVA (17MCA05)

INTRODUCTION:

Digital Libraries are being created today for diverse communities and indifferent fields e.g. education, science, culture, development, health, governance and so on. with the availability of several free digital Library software packages at the recent time, the creation and sharing of information through the digital library collections has become an attractive and feasible proposition for library and information professionals around the world . Library automation has helped to provide easy access to collections through the use of computerized library catalogue such as On-line Public Access Catalog (OPAC). Digital libraries differ significantly from the traditional libraries because they allow users to gain an on-line access to and work with the electronic versions of full text documents and their associated images. Many digital libraries also provide an access to other multi-media content like audio and video.

PROBLEM STATEMENT:

The purpose of the Library Management system is to allow for storing details of a large number of books, magazines, Journals, thesis and allow for add, search, borrow, return facilities separately to administrator/Librarian, staff and students. Different privileges are given to different types of users.

THE OBJECTIVES ARE:

1. To collect, organize & collate print & digital information & disseminate at the point of care and for future use.
2. To provide seamless access to information
3. To act as gateway to digital and electronic information
4. To develop in to a single access point library.

METHODOLOGY:

Digitization has opened up new audiences and services for libraries, and it needs to be integrated into the plans and policies of any institution to maximize its effectiveness. Digitization is a complex process with many crucial dependencies between different stages over time. Utilizing a holistic life-cycle approach for digitization initiatives will help develop sustainable and successful project.

REFERENCE:

http://www.digilib.org/collections

BANK CUSTOMER MANGEMENT SYSTEM

KEERTHI K PHIRANGI (17MCA12)

POOJASHREE.N (17MCA18)

INTRODUCTION:

Today, banks are looking beyond the transactions to the full opportunity on how to manage their customers. Accordingly, they are moving beyond managing clients as simple contacts to a whole new level of client relationship management, crafting a superior commercial client experience that gives the bank a competitive advantage and a more loyal, profitable and committed customers. Internet Banking System refers to systems that enable bank customers to access accounts and general information on bank products and services through a personal computer or other intelligent device. But most of these systems do not focus on how best to manage and keep their customer's data more secured. The chances and threats that the internet symbolizes is no longer news to the present day banking sector. No traditional bank would dare face investment analysts or new customers without an internet strategy. The main intention behind the commencement of electronic banking services is to provide the customers with an alternative that is more responsive and with less expensive options. With options more secured, customers have more control than ever. Their expectations are how safe and secured their personal information would be. They also want personal attention and highly customized products and services. This Bank Customer Management System (BCMS) aims to provide critical information for managing the bank customers more effectively, and encourages other banks that already have similar systems to move beyond transactions to better customer management approach.

PROBLEM STATEMENT:

At present most of the banking applications are yet to overcome the rapidly growing attacks on their customer private data. Issues suck fraud operating within a conventional environment. However current systems are still trying to cope with the existing institutional structure, which is really meant for usual banking system only and not managing their customer's information more effectively.

THE OBJECTIVES ARE:

1. To create a banking system that is easily via internet
2. Reduce the flow of human traffic and long queues at banks
3. Reduce the time wasted in going to banks to update personal details.
4. To develop a bank customer management system with a multi-level security measure that will restore the customers' confidence.

METHODOLOGY:

A software development methodology is a framework that is used to structure, plan, and control the process of developing an information system, this includes the pre-definition of specific deliverables and artifacts that are created and completed by a project team to develop or maintain an application. A wide variety of such frameworks have evolved over the years, each with its own recognized strengths and weakness. One software development methodology framework is not necessarily suitable for use by all projects.

REFERENCE:

http://www.ijstr.org/final-print/aug2015/Bank-Customers-Management-System.pdf

IMPLEMENTATION OF A SENSOR DATABASE SYSTEM USING A GENERIC DISSEMINATION MECHANISM

ROJA RANI K (17MCA23)

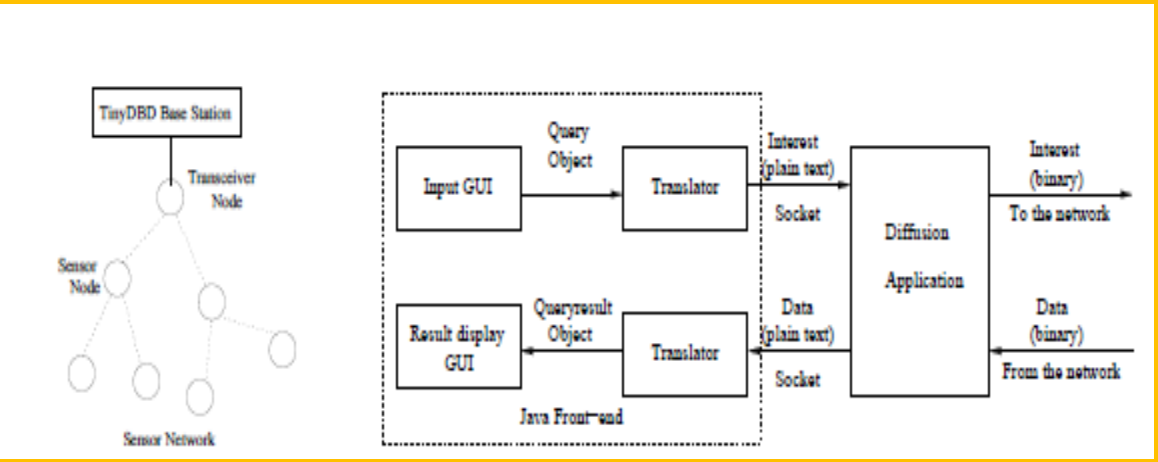
AISHWARYA M (17MCA01)

INTRODUCTION:

Prior work has proposed the use of database-like interfaces to program sensor networks. Enabling technology that uses SQL to program sensor networks makes large-scale and fine grained sensing accessible to scientists and researchers in other disciplines. Several such sensor database systems have been prototyped and some have even been deployed. TinyDB is one such system. Tiny DB allows users to query a sensor network using SQL queries. In response to the SQL queries, query processors that run in each node in the network process and aggregate streams of sensor values much like how streams are processed in a database. TinyDB developers have crafted a networking mechanism tailored specially for routing and topology maintenance in a TinyDB network. When queries are propagated to the network, a tree is formed which is used to route data back to the base station

DESIGN:

Tiny DB on Diffusion (Tiny DBD) exports an SQL interface to the end users of a sensor network Tiny DBD uses the filter architecture of Directed Diffusion to implement query processing mechanisms in the sensor nodes. It consists of three major components: (a) The base station, (b) Filter based query processing, and (c) Sensors. **The base station.** The base station is a PC node that allows a user to formulate a query using a GUI. The station injects the query into the network using the transceiver node. TinyDBD supports multiple base stations injecting a query to the network at the same time. Each base station runs its own instance of the front end and the glue code. In a TinyDBD base station there are two main components: the (1) TinyDB front-end, which communicates with (2) Diffusion-based base station code. Figure 1 shows the architecture of the base station software.



IMPLEMENTATION:

Tiny DBD was implemented using Diffusion 3.1.2 libraries in C++. The sensors and the front end make Publish and Subscribe calls. The database logic is written in a filter called DB Filter. We have tested Tiny DBD on a10-node PC104 network. Our limited experiments show that it is possible to implement a database-like system on Diffusion. During the implementation of Tiny DBD, we decided to use a TCP socket interface between the Tiny DB front end and Diffusion-based base station software rather than integrating these two pieces of software.

REFERENCES:

- P. Bonnet, J. E. Gehrke, and P. Seshadri. Towards Sensor Database Systems. In *MDM*, 2001.
- P. Buonadonna, D. Gay, J. Hellerstein, W. Hong, and S. Madden. TASK: Sensor Network in a Box. In *EWSN*, 2005.

MOBILE DATABASE

Y.HARICHANDANA (17MCA29)

TEJASWINI G (17MCA27)

INTRODUCTION:

A mobile database is a database that can be connected to by a mobile computing device over a wireless communication mobile networks. Database that is portable and physically separate from the corporate database server, It can also be defined as system with the following structural and functional properties. Distributed system with mobile connectivity.

MOBILE DATABASES:

Physically separate from the central database server Reside on mobile devices. Capable of communicating with a central databases server of other mobile clients databases server. Complete spatial mobility

Eg: smart phones and PDAs store and share data over a mobile network,

CHARACTERISTICS:

- Limited power supply.
- Limited resources.
- Mobility
- Disconnections.

CURRENT APPOROACH:

Currently most mobile application developers uses “flat files” to store application. Data “Flat Files” is a file containing records that have no structured interrelationship**.**

ADVANTAGE:

Smaller and easier to manage Greater access to modern apps and services.

DISADVANTAGES:

Application need to know the organization of the records within the file. Developers have to implement the required databases functionalities.

EXISTING MOBILE DATABASES:

Sybase SQL Anywhere: It is a relational database management system is a product of sap.

Oracle Lite: It is a single user relational database that runs on desktop, laptop, down smallest help devices.

Microsoft SQL server compact: it is optimized for an architecture where all applications share the memory pool.

SQLite: SQLite is an open source mobile database engine.

EMBEDDED DATABASES:

Tiny DB: It supports only select operations of the standard SQL. Its memory footprint is only 3kbytes, it supports only essential functionalities for sensor applications.

Pico DBMS: It supports sufficient functionalities for smart card applications, smart card applications are used for data management such as insert, delete, update and search ,footprint is 30kbytes.

APPLICATIONS:

Business: Information on customers, competitors, and marketing trends anytime.

Public sector: Army uses mobile databases technology to get current inventory info that can save.

Health: Used by physicians to store and retrieve info while making their rounds. Used by doctors and para medics to retrieve vital patient history & treatment info while attending to patients in battle field &remote accident locations.

REFERENCE:

Nori A mobile and embedded databases.in proc. SIGMOD (2007 edition) DBMS techniques for lite weight computing devices in proc. Mobide (2011 edition) Mobile application and other information in proc. Whangk y (2009 edition).

CLUSTERING IN DATA MINING

AMALASEELAN G V (17MCA02)
DIVYA G M (17MCA08)

INTRODUCTION:

Cluster is a group of objects that belong to the same class. In other words the similar object are grouped in one cluster and dissimilar are grouped in other cluster. A cluster of data objects can be treated as a one group. While doing the cluster analysis, we first partition the set of data into groups based on data similarity and then assign the label to the groups.

Clustering is the process of making group of abstract objects into classes of similar objects. The main advantage of Clustering is that, it is adaptable to changes and help single out useful features that distinguished different groups.

REQUIREMENTS OF CLUSTERING IN DATA MINING:

- **SCALIBILITY** - We need highly scalable clustering algorithms to deal with large databases. Ability to deal with different kind of attributes - Algorithms should be capable to be applied on any kind of data such as numerical data, categorical, binary data.
- **HIGH DIMENSIONALITY** - The clustering algorithm should not only be able to handle low- dimensional data but also the high dimensional space.
- **INTERPRETABILITY** - The clustering results should be interpretable, comprehensible and usable.

CLUSTERING METHODS

The clustering methods can be classified into following categories:

- **PARTIONING METHOD**-Suppose we are given a database of n objects, the partitioning method construct k partition of data. Each partition will represents a cluster and $k \leq n$.

- Each object must belong to exactly one group.
- Each group contain at least one object.

- **HIERARCHICAL METHOD**- This method create the hierarchical decomposition of the given set of data objects. We can classify Hierarchical method on basis of how the hierarchical decomposition is formed as follows:
  - Agglomerative Approach
  - Divisive Approach

- **GRID-BASED METHOD**- In this the objects together from a grid. The object space is quantized into finite number of cells that form a grid structure.

- **MODEL-BASED METHODS**- In this method a model is hypothesize for each cluster and find the best fit of data to the given model. This method also serve a way of automatically determining number of clusters based on standard statistics. It therefore yield robust clustering methods.
- **CONSTRAINT-BASED METHOD**-In this method the clustering is performed by incorporation of user or application oriented constraints. The constraint can be specified by the user or the application requirement.

REFERENCE

http://bigdata-madesimple.com/what-is-clustering-in-data-mining/

http://www.cs.put.poznan.pl/jstefanowski/sed/DM-7clusteringnew.pdf

www.cc.gatech.edu/~isbell/reading/papers/berkhin02survey.pdf

ORACLE VERSUS MS SQL SERVER

MARY HARSHITHA A (17MCA15)

SWEETY LENKA (17MCA26)

INTRODUCTION:

There are many different database management systems out there. We would have heard about Microsoft Access, Sybase, and My SQL, but the two most popular and widely used are Oracle and MS SQL Server. Although there are many similarities between the two database management systems, there are also a number of key differences. The Oracle Database is an object-relational database management system (ORDBMS). It is commonly referred to as Oracle RDBMS or simply as Oracle. Software Development Laboratories (SDL) developed the original version of the Oracle software. MS SQL Server is a relational database management system (RDBMS) developed by Microsoft. This product is built for the basic function of storing retrieving data as required by other applications. It can be run either on the same computer or on another across a network.

LANGUAGE

The most obvious difference between the two database management systems is the language they use. Although both systems use a version of Structured Query Language, MS SQL Server uses Transact SQL, or T-SQL, which is an extension of SQL originally developed by Sybase and used by Microsoft. Oracle, meanwhile, uses PL/SQL, or Procedural Language/SQL. Both languages have different syntax and capabilities. The main difference between the two languages is how they handle variables, stored procedures, and built-in functions. PL/SQL is complex and potentially more powerful, while T-SQL is much more simple and easier to use.

TRANSACTION CONTROL

Another one of the biggest differences between Oracle and MS SQL Server is transaction control. A transaction can be defined as a group of operations or tasks that should be treated as a single unit. For instance, a collection of SQL queries modifying records that all must be updated at the same time, where (for instance) a failure to update any single records among the set should result in none of the records being updated. By default, MS SQL Server will execute and commit each task individually, and it will be difficult or impossible to roll back changes if any errors are

encountered along the way. To properly group statements, the "BEGIN TRANSACTION" command is used to declare the beginning of a transaction, and either a COMMIT statement is used at the end. This COMMIT statement will write the changed data to disk, and end the transaction. Within a transaction, ROLLBACK will discard any changes made within the transaction block. When properly used with error handling, the ROLLBACK allows for some degree of protection against data corruption. After a COMMIT is issued, it is not possible to roll back any further than the COMMIT command.

Within Oracle, on the other hand, each new database connection is treated as new transaction. As queries are executed and commands are issued, changes are made only in memory and nothing is committed until an explicit COMMIT statement is given (with a few exceptions related to DDL commands, which include "implicit" commits, and are committed immediately). After the COMMIT, the next command issued essentially initiates a new transaction, and the process begins again. This provides greater flexibility and helps for error control as well, as no changes are committed to disk until the DBA explicitly issues the command to do so.

ORGANIZATION OF DATABASE OBJECTS

The last difference is how database management systems organizes database objects. MS SQL Server organizes all objects, such as tables, views, and procedures, by database names. Users are assigned to a login which is granted accesses to the specific database and its objects. Oracle, all the database objects are grouped by schemas, which are a subset collection of database objects and all the database objects are shared among all schemas and users. Even though it is all shared, each user can be limited to certain schemas and tables via roles and permissions.

REFERENCE:

www.google.com

An Introduction to Database Systems (8th Edition) [C.J. Date]

Database System Concepts Sixth Edition

HUMAN RESOURCE MANAGEMENT SYSTEM

SUSHMA M N (17MCA25)

RAKSHITHA (17MCA22)

INTRODUCTION:

Human beings are social beings and hardly ever live and work in isolation. We always plan, develop and manage our relations both consciously and unconsciously. The relations are the outcome of our actions and depend to a great extent upon our ability to manage our actions. From childhood each and every individual acquire knowledge and experience on understanding others and how to behave in each and every situations in life. Later we carry forward this learning and understanding in carrying and managing relations at our workplace. The whole context of Human Resource Management revolves around this core matter of managing relations at work place.

That predicted future is today's reality. Most managers in public- and private sector firms of all sizes would agree that people truly are the organization's most important asset. Having competent staff on the payroll does not guarantee that a firm's human resources will be a source of competitive advantage. However in order to remain competitive, to grow, and diversify an organization must ensure that its employees are qualified, placed in appropriate positions, properly trained, managed effectively, and committed to the firm's success. The goal of HRM is to maximize employees' contributions in order to achieve optimal productivity and effectiveness, while simultaneously attaining individual objectives (such as having a challenging job and obtaining recognition), and societal objectives (such as legal compliance and demonstrating social responsibility).

Human resources management (HRM) is a management function concerned with hiring, motivating and maintaining people in an organization. It focuses on people in organizations. Human resource management is designing management systems to ensure that human talent is used effectively and efficiently to accomplish organizational goals.

OBJECTIVES ARE:

1) Human capital: assisting the organization in obtaining the right number and types of employees to fulfill its strategic and operational goals

2) Developing organizational climate: helping to create a climate in which employees are encouraged to develop and utilize their skills to the fullest and to employ the skills and abilities of the workforce efficiently

3) Helping to maintain performance standards and increase productivity through effective job design; providing adequate orientation, training and development; providing performance-related feedback; and ensuring effective two-way communication.

4) Helping to establish and maintain a harmonious employer/employee relationship

5) Helping to create and maintain a safe and healthy work environment

6) Developing programs to meet the economic, psychological, and social needs of the employees and helping the organization to retain the productive employees

METHODOLOGY:

A software development methodology is a framework that is used to structure, plan, and control the process of developing an information system, this includes the pre-definition of specific deliverables and artifacts that are created and completed by a project team to develop or maintain an application. A wide variety of such frameworks have evolved over the years, each with its own recognized strengths and weakness. One software development methodology framework is not necessarily suitable for use by all projects

REFERENCE:

http://www.dphu.org>books>books_1493_0.human_resource_managment.pdf

CLOUD COMPUTING

NANDINI D R (17MCA16)

HARSHITHA M (17MCA10)

INTRODUCTION:

A cloud computing refers to delivery of computing different resources and information over the internet between different devices located at different places. It provides services to the access of data over internet to store information at another location instead of keeping data on our own hard-disk or application for our need. The use of data can be done by software and hardware resources which are managed by third parties at remote location.

CLOUD COMPUTING CONCEPT:

There are certain services and models working behind the scene making the cloud computing feasible and accessible to end users. Following are the working models for cloud computing:

DEPLOYMENT MODEL:

Deployment models define the type of access to the cloud. Cloud can have any of the four types of access: public, private, hybrid and community.

Public cloud: the public cloud allows system and services to be easily accessible to the general public. Public cloud may be less secure because of its openness. Eg: e-mail.

Private cloud: the private cloud allows system and services to be easily accessible to the within the organization.it offers increased security because of its private nature.

Community cloud: the community cloud allows system and services to be accessible by group of organization.

Hybrid cloud: the hybrid cloud is mixture of public and private cloud. The critical activities are performed using private cloud while the non-critical activities are performed using public cloud.

SERVICE MODEL:

1. **Infrastructure as a Service (IaaS):** IaaS is a form of cloud computing which provides a virtualized computing resources over the internet. IaaS provides the underlying operating systems, security, networking, and servers for developing applications, services, and for deploying development tools, databases, and soon on.

2. **Platform as a Service (PaaS):** PaaS provides a platform for customers to develop, run, and manage applications without knowing the complexity of building and maintenance of the infrastructure in the app.

3. **Software as a Service (SaaS):** SaaS is based on the software licensing and delivery model. It is also referred as "on-demand software". It is accessed by users in the web browser by using thin client.

ADVANTAGE:

- Lower computer costs and performance is high.
- Unlimited storage capacity and provide universal document access.
- Increased data reliability and device independence.

DISADVANTAGE:

- Requires a constant internet connection and sometimes can be slow.
- Does not work well with low-speed connections.
- Stored data is lost and might not be secure.

REFERENCE:

Cloud Computing: From Beginning to End, by Ray J. Rafaels

Cloud computing: concepts, technology and architecture, by Thomas Erl

GEOGRAPHIC INFORMATION SYSTEM

SANJANA K R (17MCA24)

A **geographic information system** (**GIS**) is a system designed to capture, store, manipulate, analyze, manage, and present spatialor geographic data. GIS applications are tools that allow users to create interactive queries (user-created searches), analyze spatial information, edit data in maps, and present the results of all these operations. Geographic information science is the science underlying geographic concepts, applications, and systems. The first known use of the term "geographic information system" was by Roger Tomlinson in the year 1968 in his paper "A Geographic Information System for Regional Planning". Tomlinson is also acknowledged as the "father of GIS"

GIS TECHNIQUES AND TECHNOLOGY:

Modern GIS technologies use digital information, for which various digitized data creation methods are used. The most common method of data creation is digitization, where a hard copy map or survey plan is transferred into a digital medium through the use of a CAD program, and geo-referencing capabilities. With the wide availability of ortho-rectified imagery (from satellites, aircraft, Helikites and UAVs), heads-up digitizing is becoming the main avenue through which geographic data is extracted. Heads-up digitizing involves the tracing of geographic data directly on top of the aerial imagery instead of by the traditional method of tracing the geographic form on a separate digitizing tablet (heads-down digitizing).

APPLICATIONS:

The implementation of GIS is often used by city purpose or application requirements. GIS provides a platform to update geographical data without wasting time to visit the field and update a database manually. GIS when integrated with other powerful enterprise solutions like SAP. Many disciplines can be benefited from GIS technology. . An active GIS market has resulted in lower costs and continual improvements in the hardware and software components of GIS, and usage in the fields of science, government, business, and industry, GIS also used in location based services, which allows GPS-enabled mobile devices to display their location.

OPEN GEOSPATIAL CONSORTIUM STANDARDS:

The Open Geospatial Consortium (OGC) is an international industry consortium of 384 companies, government agencies, universities, and individuals participating in a consensus process to develop publicly available geoprocessing specifications. GIS products are divided into 2 categories by OGC Open Geospatial Consortium protocols include Web Map Service, and Web Feature Service. Open GIS Specifications support interoperable solutions that "geo-enable" the Web, wireless and location-based services, and mainstream IT, and empower technology developers to make complex spatial information and services accessible and useful with all kinds of applications.

WEB MAPPING:

In recent years it has been very much easy to access mapping software such as google maps and Bing maps and also free open source map. These services give the public access to huge amounts of geographic data; perceived by many users to be as trustworthy and usable as professional information. Some of them, like Google Maps and Open Layers, expose an application programming interface (API) that enable users to create custom applications. These toolkits commonly offer street maps, aerial/satellite imagery, geocoding, searches, and routing functionality. These mashup projects have been proven to provide a high level of value and benefit to end users outside that possible through traditional geographic information.

ADDITION OF THE DIMENSIONS OF TIME:

The condition of the Earth's surface, atmosphere, and subsurface can be examined by storing satellite information into a GIS. GIS technology gives researchers to examine the variations in Earth processes over days, months, and years as an example, the changes in vegetation vigor through a growing season can be done to determine when drought was most extensive in a particular region. The resulting graphic represents a rough measure of plant health. Working with two variables over time would then give the ability to researchers to detect regional differences in the lag between a decline in rainfall and its effect on vegetation.

REFERENCES:

https://en.wikipedia.org

www.nationalgeographic.org

DISTRIBUTED DATABASE

KHUSHABOO KUMARI (17MCA13)

PUJA KUMARI (17MCA20)

- A distributed database is a collection of multiple interconnected, which are spread physically across various locations that communicate via a computer network.
- A distributed database (DDB) is a collection of multiple logically interrelated database distributed over a computer network.
- A distributed database management system (D-DBMS) is the software that manage the DDB and provides an access mechanism that makes this distributed transparent to the users.
- Distributed database system (DDBS) = DDB + D-DBMS.

DISTRIBUTED DATABASE MANAGEMENT SYSTEM:-

A distributed database management system (DDBMS) is a centralized software system that manages a distributed database in a manner as if it were all stored in a single location.

FEATURES OF DISTRIBUTED DATABASE SYSTEM:-

It is used to create, retrieve, update and delete distributed databases.

- It synchronizes the database periodically and provides access mechanisms by the virtue of which the distribution becomes transparent to the users.
- It ensures that the data modified at any site is universally updated.
- It is used in application areas where large volumes of data are processed and accessed by numerous users simultaneously.
- It is designed for heterogeneous database platforms.
- It maintains confidentiality and data integrity of the databases.

ADVANTAGES OF DISTRIBUTED DATABASES:

Modular Development − If the system needs to be expanded to new locations or new units, in centralized database systems, the action requires substantial efforts and disruption in the

existing functioning. However, in distributed databases, the work simply requires adding new computers and local data to the new site and finally connecting them to the distributed system, with no interruption in current functions.

More Reliable − In case of database failures, the total system of centralized databases comes to a halt. However, in distributed systems, when a component fails, the functioning of the system continues may be at a reduced performance. Hence DDBMS is more reliable.

Better Response − If data is distributed in an efficient manner, then user requests can be met from local data itself, thus providing faster response. On the other hand, in centralized systems, all queries have to pass through the central computer for processing, which increases the response time.

Lower Communication Cost − In distributed database systems, if data is located locally where it is mostly used, then the communication costs for data manipulation can be minimized. This is not feasible in centralized systems.

ADVERSITIES OF DISTRIBUTED DATABASES:-

- **Need for complex and expensive software** − DDBMS demands complex and often expensive software to provide data transparency and co-ordination across the several sites.
- **Processing overhead** − Even simple operations may require a large number of communications and additional calculations to provide uniformity in data across the sites.
- **Data integrity** − the need for updating data in multiple sites pose problems of data integrity.
- **Overheads for improper data distribution** − Responsiveness of queries is largely dependent upon proper data distribution. Improper data distribution often leads to very slow response to user requests.

REFERENCE:

https://www.tutorialspoint.com/distributed_dbms/distributed_dbms_databases.htm

MULTIMEDIA DATABASE MANAGEMENT SYSTEM

BHAVYASHREE R (17MCA21)

AMALI SUNITHA P (17MCA03)

INTRODUCTION:

In this era of computers, multimedia has emerged into our day to day life. Multimedia data typically means digital images, audio, video, animation and graphics together with data. Which can capture and display multimedia data. In today's world not only personal data but professional data is digitized. Data and results of experiments are stored in digitized form which make backup copies and provide access to other professional's easier.

A **Multimedia Database Management System** (**MMDBMS**) is a framework that manages different types of data potentially represented in a wide diversity of formation a wide array of media sources. It provides support for multimedia data types and facilitate for creation, storage, access, query and control of a multimedia database. The basic concentration is on how to retrieve the stored data images effectively in less time as compared to other ways of retrieval. The ultimate objective is how to access multimedia data effectively. With respect to access multimedia data can be classified into active objects i.e. those objects which participate in retrieval process and passive objects .In order to retrieve data from database system queries are needed.

GENERIC ARCHITECTURE OF MULTIMEDIA DATABASE:

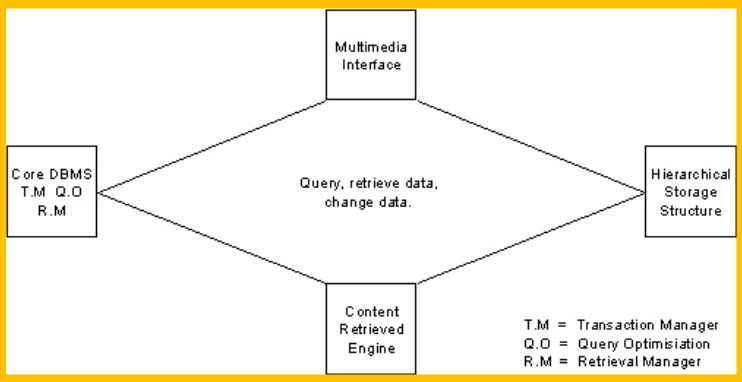
Multimedia Databases require all the basic attributes of a database management system such as a transaction manager, query optimizer, recovery manager etc. as well as special storage structures and specialized search and querying modules.

REQUIREMENTS:

- **Integration**
  - Data items do not need to be duplicated for different programs invocations
- **Data independence**
  - Separate the database and the management from the application programs

- **Concurrency control**
  - Allows concurrent transactions
- **Persistence**
  - Data objects can be saved and re-used by different transactions and program invocations
- **Privacy**
  - Access and authorization control
- **Integrity control**
  - Ensures database consistency between transactions
- **Recovery**
  - Failures of transactions should not affect the persistent data storage
- **Query support**
  - Allows easy querying of multimedia data

However, information retrieval and document imaging systems require searching the content of documents. This ability can be generalized to still images, audio and video.



LATEST TECHNOLOGY OF MMDMS:

In a networked Management System such as video conferencing, real-time image communications is the system success. Also, future research directories are indicated to motivate the researchers in the multimedia application in Business.

REFERENCES:

http://aircconline.com/ijma/V3N1/3111ijma01.pdf

https://www.researchgate.net/publication/3814848_Issues_in_multimedia_database_management

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.353.5639&rep=rep1&type=pdf

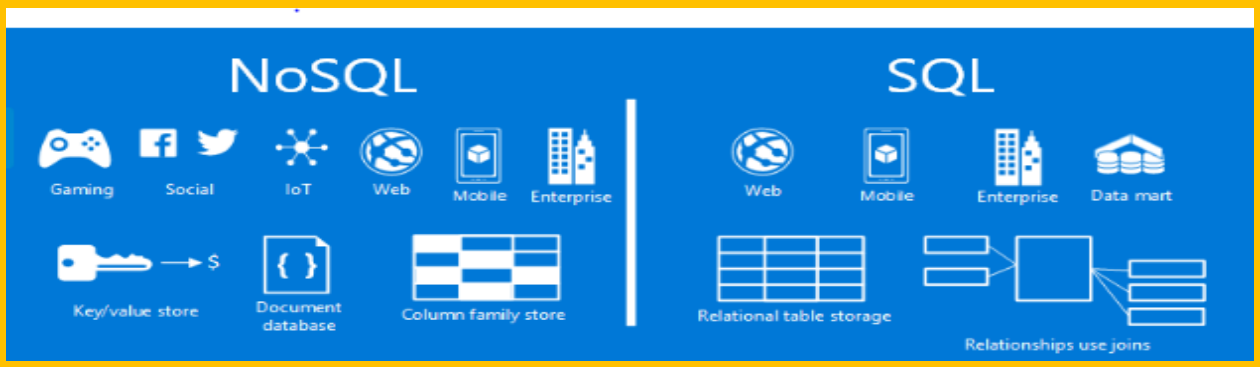
SQL v/s NOSQL

ANNAPURNA PAILA (17MCA04)

JOTHIS MARIA (17MCA11)

INTRODUCTION:

In the world of database technology, there are two main types of databases: SQL and NoSQL—or, relational databases and non-relational databases. The difference speaks to how they're built, the type of information they store, and how they store it. Relational databases are structured, like phone books that store phone numbers and addresses. Non-relational databases are document-oriented and distributed, like file folders that hold everything from a person's address and phone number to their Facebook likes and online shopping preferences. We call them SQL and NoSQL, referring to whether or not they're written solely in structured query language (SQL). In this journal.



What Is SQL?

SQL is a language that facilitates communication with relational database management systems, most of which have their exclusive proprietary extensions. SQL can do everything from accessing and manipulating databases to inserting records and creating views.

- **MySQL**—the most popular open-source database, excellent for CMS sites and blogs.
- **Oracle**—an object-relational DBMS written in the C++ language. If you have the budget, this is a full-service option with great customer service and reliability. Oracle has also released an Oracle NoSQL database.

- **IMB DB2**—a family of database server products from IBM that are built to handle advanced "big data" analytics.
- **Sybase**—a relational model database server product for businesses primarily used on the Unix OS, which was the first enterprise-level DBMS for Linux.

Microsoft-Developed RDBMS For Enterprise-Level Databases That Supports Both Sql And NOSQL Architectures.

- **Microsoft Azure**—a cloud computing platform that supports any operating system, and lets you store, compute, and scale data in one place. A recent survey even put it ahead of Amazon Web Services and Google Cloud Storage for corporate data storage.
- **MariaDB**—an enhanced, drop-in version of MySQL.
- **PostgreSQL—**an enterprise-level, object-relational DBMS that uses procedural languages like Perl and Python, in addition to SQL-level code.

What Is NoSQL? NoSQL encompasses a wide range of database technologies that are designed to cater to the demands of modern apps. NoSQL systems make it easy to deploy and store a wide range of data types, and they excel in: If your data requirements aren't clear at the outset or if you're dealing with massive amounts of unstructured data, you may not have the luxury of developing a relational database with clearly defined schema

The Key Differences between NoSQL and SQL

	SQL	**NOSQL**
Storage	NoSQL encompasses a host of database types ranging from graph and key-value to document and columnar.	Data is typically stored in a relational model where columns contain data points and rows comprise of all the information

		concerning a single entity.
Flexibility	Since schemas are dynamic in nature, information can be updated on the fly.	In SQL, every record conforms to a predefined schema where the columns must be determined and locked before the data can be entered and it cannot be amended later without going offline.
ACID Compliance	NoSQL emphasizes performance over data integrity and most NoSQL systems compromise on ACID compliance for performance, so organizations use NoSQL for data types not impacted by consistency.	SQL databases default to enabling ACID compliance though most offer options to favor performance over data integrity for some operations (e.g., asynchronous replication between sites can risk data loss during failure).

REFERENCE:

https://int.search.tb.ask.com

https://upgrad.com/data-science