CHAPTER – IV

SYSTEM DESIGN

The system design and development is explained in details in this chapter it’s important to keep track of the various features and modules of the system designed

Figure 4.1 Various Components of the Research Project

The above figure 4.1 explains all the components which are connected and are requesting information from the central repository this information is required by the users and various reports are generated these reports serve as a means of system design and development module the stake holders connect to the central repository and access the data as an when required by the modules [Soe-Tser Yuan,2009]. The project manager connects to the data warehouse and collects the data about all the ongoing projects and views the status of each project the stake holders assigned to the projects can also be seen in the project report that is
generated at the project managers end this system design and development is more connected and collaborative in nature hence it’s easy to access the data from the various repositories and collect information at various stages of the project.

The system is designed and developed to cater the needs of various stakeholders of the software engineering project all the data which is required by the various stakeholders is stored in the repository information management is important and its necessary to manage all the requirements at a central location without any information getting lost [Eunchang Lee,2009]. The data warehouse is a secured place with good bandwidth connection and collaboration of the modules today we need to be more connected central resources with high data availability with efficiency.

![Diagram of stakeholders](image)

**Information Mining**

Figure 4.2 Information Extraction by various stakeholders

The figure 4.2 above shows how the stakeholders connect to the data warehouse and extract the information from the various sources and phases of the project management all the information extraction is done using the data mining algorithms the K-means algorithm is used to extract the relevant information the data mining is performed on the data warehouse information [Tzu-Chi Huang,2012]. The actual data mining tasks such as data record groups (cluster data analysis), unusual records (anomaly information detection) and dependencies (association rule mining), as large quantities of data to extract previously unknown interesting patterns of the automatic or semi-automatic analysis. This is usually as a spatial
index is using database technologies. This pattern can be seen as a kind of summary of the input data, and machine learning and predictive analysis, or can use more analysis. For example, data mining and decision support system by step to obtain a more accurate prediction results can be used; the data is able to identify several groups[Isabel Ramos Roman,2008]. Data collection, data preparation, interpretation and reporting of the results of the two data-mining step, but as an additional step in the process is not followed for extraction and management of the information. Data mining may arise as a result of a specific problem or situation analysis uses data from the past data. Data mining analysis is used to store the data that is stored in data warehouses for data analysis work. Accurate data for the management of production, from all parts of the business. Marketing strategy for their product managers to decide on the use Mining information. They compare and contrast between competitors can use the data. Data mining, selling to promote the new product, or a company - not a value added product that can be used to interpret its data in real-time analysis [Ashley Chonka,2011].

4.1 Pre-processing

Before data mining algorithms can be used, a target data set must be assembled. Data mining that data to uncover patterns actually present, the target data set to be created within an acceptable time limit, while remaining compact enough to be included in the pattern must be sufficiently large. Data mart or data warehouse is a common source of information. Before the pre-processing of data mining to analyze multivariate data sets are required. The target set is then cleaned [Gilles Fedak,2009]. Data cleaning noise and the accompanying observations with missing data are removed. The duplication of the information needs to be removed from the data so that a single and updated copy of the project is maintained. Most of the software engineering project today needs to be analyzed and developed in system utility type [Luis M. Alvarez Sabucedo,2009].

The final step in the data mining and knowledge discovery from data produced by the algorithm is to determine the pattern of a large data set. All patterns found by the data mining algorithms are not necessarily valid. Data mining algorithms, a common data set is not available, it is common to find patterns in the training set. This is called over fitting. To avoid this, evaluation of the data mining algorithm was not trained on the use of a test data set. The learned patterns are applied to the test set, and the resulting output is compared with the desired output. For example, "legitimate" emails from the "spam" data-mining algorithm
trying to distinguish a simple e-mail will be trained on the training set [Christian Vecchiola, 2012]. Once trained, the learned patterns that had not been trained to set up e-mail on which the test will be applied to. The accuracy of the pattern of e-mails and how they can be measured correctly classified. A number of statistical methods such as curve algorithm can be used to assess by the users of the system.

4.2 Result Validation

The information, which is stored on the data warehouse, should be validated with the exact data otherwise there can be duplication of the information and records in the system can be over written by the new data published the central repository of the system it’s important to cater to the needs of the information validation before uploading any information on the system repository of over writing the previous data the backup copy of the information should be maintained with all the records synchronized and a true copy of the information should also be maintained on the data warehouse this is important for comparison of data with various stages of the project management system [Andrea Bosin, 2011]. Data mining to reveal hidden patterns and trends in data warehouse database as a collection of static data analysis of historical business activities are. Data mining software to detect previously unknown strategic business information to help you sift through large amounts of data to use advanced pattern recognition algorithms. Businesses prevent customer attrition and new customers with more accuracy, cross-sell to existing customers and to profile customers, to find the root cause of problems, identify new production space analyzing the market, including what are examples of the use of data mining. Today, companies are collecting an explosion at the raw data. For example, Walmart every 20 million points-of-sale transactions and processes [Pablo Bermell, 2012]. This information is stored in a central database, but to analyze the data mining software would be useless without some kind. Walmart data mining techniques to analyze their point-of-sale data to determine trends in the sales and marketing campaign development and will be able to more accurately predict customer loyalty [Sriram Krishnan, 2009]. We credit card, a store loyalty card, or fill out warranty card information every time you use our purchasing power is being collected about behavior. Many people, such as Google, Facebook, and Amazon, as distributing companies, find the amount of information collected about us and are concerned about privacy. Harmful, or unwanted can be used for our personal information is possible, although it is also used to make our lives better. They recommend ways to safe and dangerous road conditions can warn drivers about so, for example, Ford and Audi hopes to collect
information about a customer driving patterns of the system that will collect the data from the data warehouse [Miguel Contreras, 2008].

Most of the companies today have sophisticated application tools, which extract data from the central repositories and dump this in the information exchanges of the system development tools [I-Chen Wu, 2012]. The key components is to validate the users inputs and the system design and development problems Today most of the tools user for data mining do not validate the data after extraction hence it’s difficult to understand the information flow how it works and how the users can collect data from the sources of the information. Today we have most of the companies collecting data from various viewpoints hence it’s important to get the information in a central place [Mani Golparvar-Fard, 2011]. Customer relationship management applications, data mining can contribute significantly to the bottom line. In addition, modern applications that can be used to automate. Data mining, The results have been determined, this "sophisticated application" either automatically send an e-mail or regular mail. Finally, many people without offering to take action in cases where the "uplift modeling" is offering a great increase in the response of the people to determine what can be used [QiuZhi, 2011]. Modeling the Advancement of persuadable people, mailings and offers enables marketers to focus, and will buy a product without offering to send offer to people who do not. Automatic clustering of data segments or groups within a customer data set can be used to find. Rather than using a model to predict how many customers will churn, a business model unique to each region and customer type can be built. Instead of sending an offer to all the people who are likely to churn, then, it may only want to send offers to loyal customers [Rosario Girardi, 2005]. Finally, business customers are going to be profitable over time certain window to determine what you can, and only those that are likely to be profitable to send offers. In order to maintain this quantity of models, they manage model versions and move to automated data mining is needed. The System Designed Models. An integrated circuit product line with the relevant data mining for example, "Very Large Scale Integrated Test Mining Test Data to Optimize." [NuriaForcada, 2007] This paper is described in the paper, die-level functional test, data mining and decision analysis to the problem of application are described. A probabilistic model to fit experimental failure patterns mining historical die test data to demonstrate the ability to apply the system is mentioned [Liu Peiyu, 2011]. This pattern is used to determine the Real-time, the next and when to stop testing to verify die. This system has the potential to improve profits on mature IC products, based on experiments with historical test data is shown. Other implications of the
semiconductor manufacturing process, on-line monitoring using data mining can be extremely effective [OutiRaiha, 2010].

4. 3 Result Authentication

The system generated result after mining of the information should be checked for testing the most of the system software require the system design and development modules which collaborate with the functionalities of the software project most of the data warehouses collect RAW data from various sources this RAW data needs to be analyzed and accessed from various system repositories [Li Chen, 2012]. Mining of information is also crucial step and the most important is Authentication of the information, which is stored in the central data base repository.

The authenticated information should be stored in different table and sections in a data warehouse so that the stakeholders will have access to the authenticated information only and no duplication of the information should be maintained [Liu Jingkuang, 2011]. The key factor concerning the information integration and exchange is important it’s important to concentrate on the global data challenges which are faced while mining the information. Most of the stack holders connect to the central repository and they get the data but there should be tools which every the data is original and no one tampered with the data when it was being transmitted from one end to the other [Gaofeng Zhang, 2012].

The information exchange from one end to the other should be secured and also encrypted so that legality of the data can be maintained and results, which are obtained from this data, are also original. Hence no flaw in the system is detected. Security of the data while extracting the RAW data from the data warehouse is also important [Javier Espadas, 2013].

4. 4 Function Point Analysis

This system calculated the function point of the software engineering project Identified and classified in a type of work that is done, then the complexity of the assessment and action points are assigned number. Such as a query to check for user input or data entry as the end user business function maps each of the functional user requirements [Jianxin Li, "CyberLiveApp, 2013]. The task of measuring points in user-friendly requirements easily map, but also the tools that are needed to implement
the built-in functions, to hide the trend, because this is an important difference, however, Changing the size of the result is an International Standard Organization recognized FSM method consists of algorithmic complexity. Recently implemented in several commercial software products to deal with the perceived weaknesses of different approaches are proposed [IoanaRus,2002]. These (and other vulnerabilities), designed to make the functional side measurement method based Albrecht differences included for calculating the system function points. Most of the software engineering project today requires the function point analysis all the function point analysis is done by the project management software [Martin Host,2007].

The function point analysis helps in calculating the system lines of codes and the man hours required for executing such kind of project with the help of function point calculation its easily understood now to manage and maintain the system design and development it’s also helps in determining the system requirement that are necessary for the project to get completed with number of man hours required and the resources required by the software engineering project. It helps in keeping track of the resource pools of the software project and manages the system design and development with all the functions and system integrated modules this system design helps to collect the data from various sources in the data warehouse [Ali Khalili,2008]. The overall design structure of the system consists of the following components, which are incorporated, in the overall design of the research software.

1. Data Warehouse

2. Data Mining Algorithm

3. User Interface

4. Extraction of the data from the warehouse

5. Display of the relevant information
These Modules are individual Modules, which communicate with each other as an when required by the users it’s important to keep track of the functionalities of the system designed. The Data warehouse will collect data from all the phases of software engineering project most of the data will be stored as RAW data in the data warehouse. The information about the projects will be needed by the various stake holders of the software engineering project to extract that data from the data warehouse the K-Means Algorithm will be used the data mining algorithm to extract the data from the various phases of the project management [Sajid Ibrahim Hashmi, 2011]. Most of the data stored in the data warehouse is modularized it’s important to extract the exact modules data and generate the reports.

The user interface designed is more intuitive and most of the stakeholders need less training since they already have an interface of the software keeping that in mind help is also engineering project. Extract and compilation of the data takes place in the central repository.

Information about the relevant module is stored in the data warehouse. The above figure 4.3 shows the overall system design and modularization of information at each stage of the software engineering project. The information flows from the data warehouse to the data-
mining module, which uses the K-Means algorithms to extract the relevant data from the data warehouse. The extracted information is then displayed on the user interface screen where the users can view the compiled result of the information, which they have requested.

4.5 Code Complexity Calculation

The code complexity module helps in determining the number of lines of code that was used while designing the software project and also the man hours and the resource persons required and the time duration taken by the developers to complete this research or any software engineering project or the module in the project [FarzanehMahdian2010 and Zhiqiang SHU,2008]. The Data Mining and Warehousing - PMS Prerequisite Phase organization established and documented policies and software engineering, software engineering, software project manager, the company must follow all the rules that are required to have been appointed. Data Mining and Warehousing - PMS prerequisites to achieving the overall project development (ie, the organization's Program Management Plan (OPMP) or Software Development Plan (SDP) and focus on the software project planning (spp) kPa are an integral part of the main methods of defining it in. the Risk Management plan or SDP OPMP adequate resources allocated to development and software project tracking and monitoring activities referred to in the complex.

Manageability and scale of the project and the scope of the problems plaguing the two is bad. These problems can only be costly; they can be fatal to your ultimate business objectives. , Definition, purpose, and scope to improve the technologies you need to take the uncertainty out of the competition. The Data Mining and Warehousing - PMS solution for you and your team more real-time understanding of the project requirements, and the scope of the inevitable changes that will transcend the boundaries of distributed teams. The visibility of each member of the team managed to control high productivity and development process that leads to the idea to create an alignment, from the beginning of the project allows for the sharing of expectations. The Data Mining and Warehousing - PMS that it is standard for application lifecycle to provide interoperability in the future without abandoning the past allows you to move. Many companies today Globalization, around the clock, nonstop design and development is very good to be true. Diverse groups, different agendas, missed opportunities, communications, and so far does not sound that projects fail: the reality is that because it is a specific point. So how you around campus or around the world to work in concert, the team focused on the distribution and the end goal is not to multiply? First, you change and configuration needs and help you maintain control of the application lifecycle in
the development of a management system that is designed to give you visibility from the integrated environment is needed [Shinichiro Kibe, 2012].

Data Mining and Warehousing - with PMS, you and your dramatic acting and communication between distributed and remote teams can improve. You can exaggerate your application development lifecycle visibility and control through the development of management on stage, you better control over the entire development process.

The product of the project / product development at different levels and different people involved in the software development will be used to collect information from. The project / product provides multi-level monitoring, the entire software development life cycle, the client and the client are working on a software development team working on both the client and developer documentation by the interaction will monitor the project / product. Company for the development of the project / product who is the client can view the documentation. Client project / product is strictly controlled different parts of the client throughout the day is to get a clear picture as their project-related per week, they can view the details of manpower and how many people are working on their project, the project / product has been allocated to the person’s condition can be monitored. This product is a software development project as an opportunity to provide assistance to industry and other small-scale and large-scale industries can be extended. Project tracking software to ensure the success of the project is an important activity. Based on quantifiable data manager must make decisions on the direction of the project. This document is both a software manager to monitor and provide the necessary information needed for the success of software projects can be an important decision that represents the processes and forms of instruction. In addition, any project is to be developed in the project to go to the client in the client’s specific Briefing will work on their projects, which can monitor the progress of the day.

The product proposal from the client to the client’s completion and delivery of the product by the company will handle all activities. will be able to communicate with the evidence, viewed by the client on the client side. The client will know the exact time and duration of manpower employed in its product development spend by each job. It is developing a product that will not be there to hide any information from the company so that project will be completed and be more secured.
4. 6 UML Diagrams

4.6.1 USECASE DIAGRAM
Figure 4.4 Use Case from Project Manager’s perspective
Figure 4.5 Use Case from Client’s perspective
Figure 4.6 Use Case from Developers perspective
4.6.2 ACTIVITY DIAGRAM

Figure 4.7 Activity Diagram

* Module functionalities are accessed via Web service(s)
NOTE:
The sequence diagram remains the same for other activities, except the screen (web page, its functionality) and data passing through the pages, web services and SQL Server.
4.6.4 DEPLOYMENT DIAGRAM

Figure 4.9 Deployment Diagram