Chapter3

3. Technologies Adopted for the Research

3.1 Introduction

In chapter 2 the review of other related studies and their importance as well as the improvements are described in a comprehensive manner. This chapter describes the technologies identified to conduct the research. How the questionnaire is developed and the data mining technology is selected for the pattern analysis. The data warehouse technology and web application design technologies are described here.

3.2 Data Collection

3.2.1 The Questionnaire Method

Everyone uses information to make decision about the future. If our information is accurate we have a probability of making a good decision. So better information is usually leads to better decision. Therefore crucial parts of a good research concerned with making sure that the questionnaire design address the needs of research. When the design of the questionnaire following features are taken into account

Decide the information required

The first step to decide what information is needed to know from the respondent in order to meet the survey objective.

Define the target respondent

Define the population where the sample data to be collected

Choose the methods(s) of reaching your target respondent

Choose an appropriate method or methods to reach the respondent in order to collect the data effectively and efficiently. As I have selected random sample and also planned to collect the data through meter readers, paper based questionnaire form has been selected for this purpose

Decide the question content

The questionnaire was based on quantitative and qualitative measures of electricity consumption about households in domestic sector. The questionnaire was designed with the objective of obtaining behavioral patterns of different groups of households that were going to be studied as a whole. These questionnaires have had different structures and contained different questions depending on the type of household. The qualitative as well as quantitative questions are added for that.

The questionnaire is basically focused on the consumption pattern and the knowledge of the households about the energy conservation habits.

Distribution of questionnaire to consumers

The questionnaire is printed and distributed through meter readers once they visit houses for meter reading for the relevant month

3.2.2 Data Collection

Gathered answers for the questionnaire from the documents are entered to Excel sheet by entering individual responses.

3.3 Technology Adapted for data warehousing

To retrieve patterns of electricity usage of domestic consumers, there should be a data repository existing. For the data warehousing SQL Server 2008 R2 is identified as the most appropriate technology for data warehousing.

SQL Server 2008 R2

SQL Server Architecture

SQL Server 2008 R2 is a complete set of enterprise-ready technologies and tools for managing databases among wide variety of networks. The features of the SQL Server 2008 R2 is as follows

- More than a simple relational database management system
- Manage organizational data
- Derive value from information within data
- High performance

- High Availability
- Secure
- Scalable

3.4 Data Mining Approach

SQL server is not a database engine itself, it is having relational and business intelligence components. According to the thesis I had to follow data mining approach, I have gone through different technologies to identify the best suited technology for data mining. According to the study I have found that SQL Server Analysis Services is the most suitable, robust and secure method for implementing data mining applications. Analysis Services provides an integrated platform for solutions that incorporate data mining. You can use either relational or cube data to create business intelligence solutions with predictive analytics[11].

Benefits of SQL Server Analysis Services

SQL Server Analysis Services is using well defined and researched principals to discover patterns in your data. This helps to make intelligent decisions about complex scenarios and problems. There are different kinds of data mining algorithms existing in Analysis Services. By Applying those mining algorithms to your data you can come up with forecasting trends, identify patterns, create rules and recommendations, doing classifications to understand the complex problems in a very advanced manner.
In SQL Server 2008, data mining is powerful, accessible, and integrated with the tools that many people prefer to use for analysis and reporting.[12]

Analysis Services – Data Mining Architecture

Following architecture diagram shows the components of Analysis Services architecture for the data mining.[13]

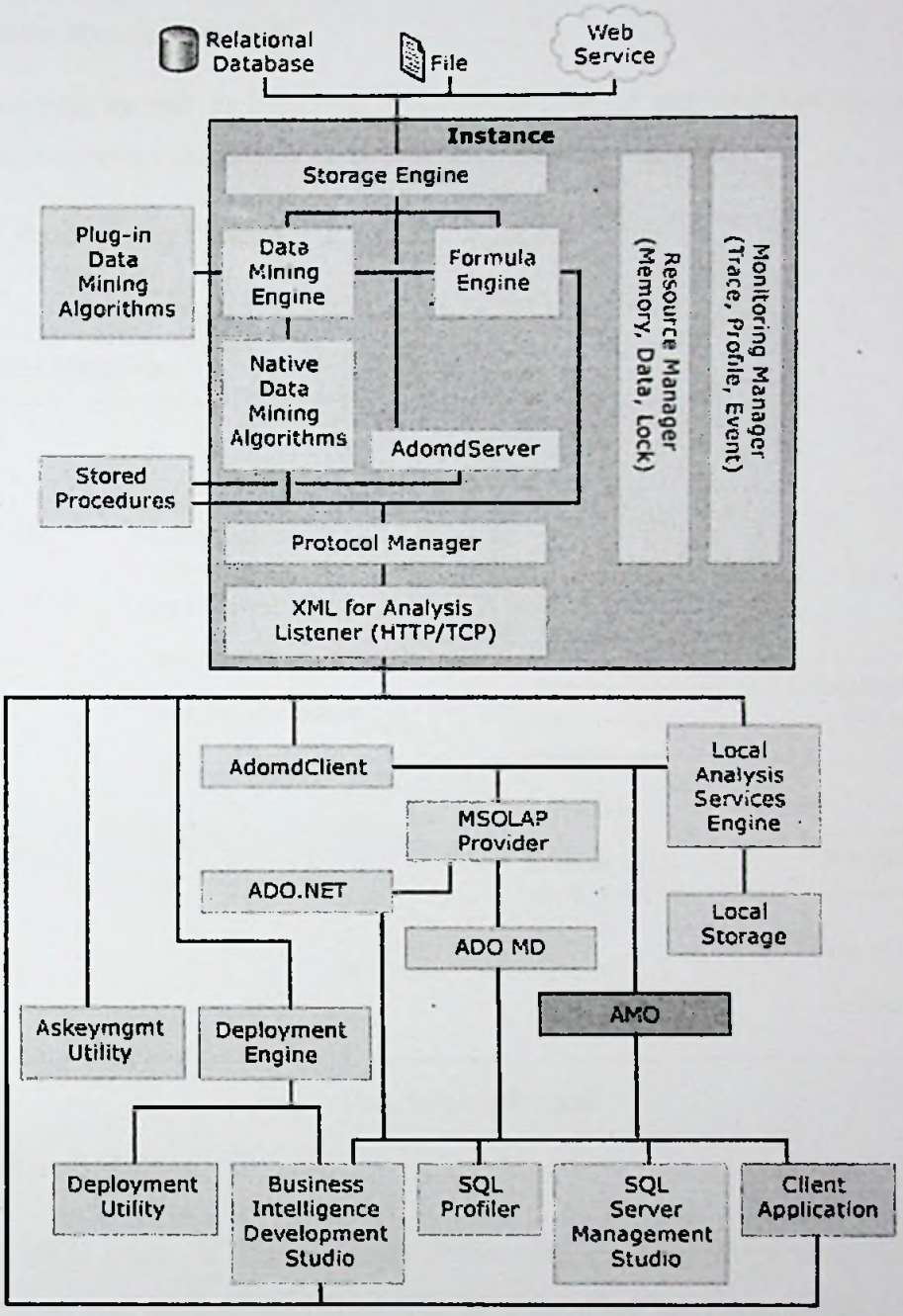

Figure 6 : Architectural Diagram of SQL Server Analysis Services Components

3.5 Data Mining Approach

Data mining technology is adopted to identify the pattern and some statistical analysis to come up with a conclusion about the hypothesis development.

3.5.1 Data Mining Process

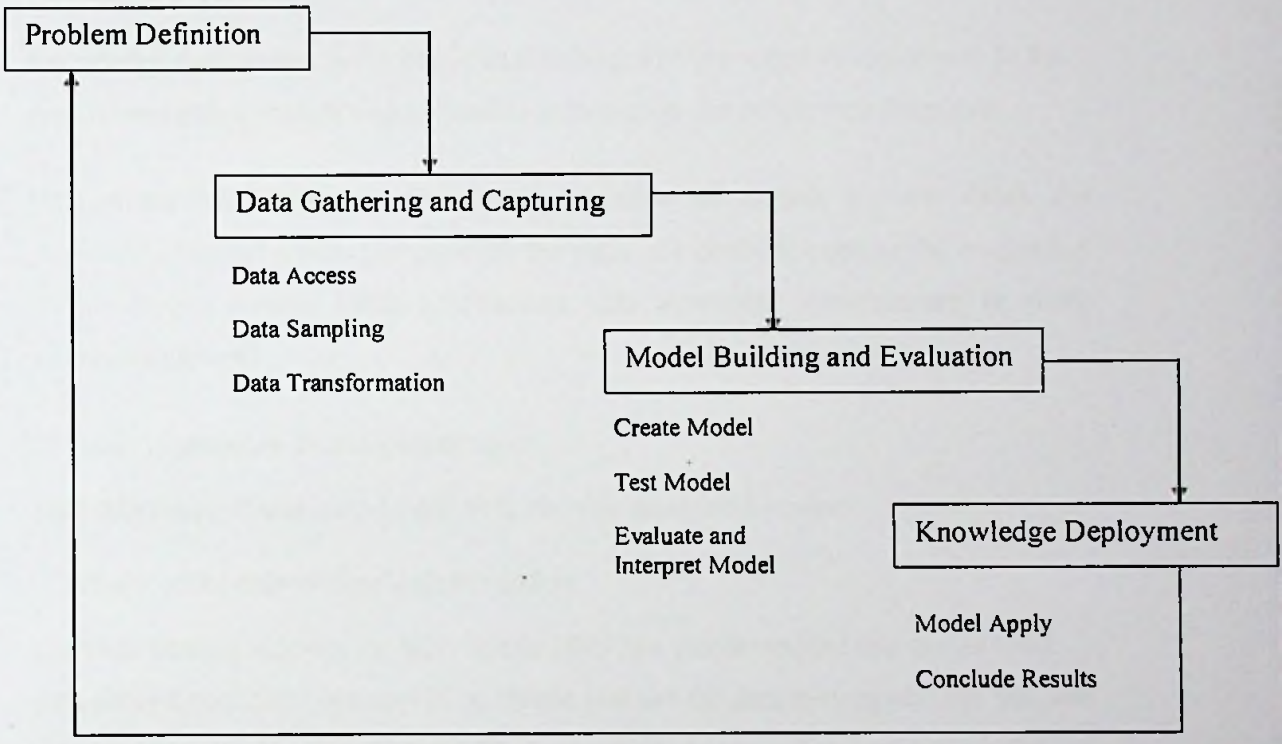

Figure 7: Data Mining Process

Problem Definition

This initial phase of a data mining project focuses on understanding the project objectives and requirements. Once you have specified the project from a business perspective, you can formulate it as a data mining problem and develop a preliminary implementation plan.

Data Gathering and Preparation

The data understanding phase involves data collection and exploration. As you take a closer look at the data, you can determine how well it addresses the business problem. You might decide to remove some of the data or add additional data. This is also the time to identify data quality problems and to scan for patterns in the data.

Model Building and Evaluation

In this phase, you select and apply various modeling techniques and calibrate the parameters to optimal values. If the algorithm requires data transformations, you will need to step back to the previous phase to implement.

Knowledge Deployment

Knowledge deployment is the use of data mining within a target environment. In the deployment phase, insight and actionable information can be derived from data.

Deployment can involve scoring (the application of models to new data), the extraction of model details (for example the rules of a decision tree), or the integration of data mining models within applications, data warehouse infrastructure, or query and reporting tools.

3.6 The Application Development tools

3.6.1 Microsoft Excel Add-In for SQL Server Analysis Services

Overview of the data mining with the add-in

The Data Mining Add-ins for SQL Server 2008 is a free download that can be used with either Excel 2007 or Excel 2010. When you use the data mining add-ins, you can connect to an existing instance of SQL Server 2008 Analysis Services and use the data mining algorithms and services provided by that server to perform data mining on the data in your Excel workbook and other supported data sources.[14]

The Data Mining Add-ins contain two sets of tools: the Table Analysis tools, which let you perform analysis by using wizards and your data in Excel, and the Data Mining Client for Excel, which provides an easy-to-user interface for building data mining models.[14]

This client tool is very easy to understand and work with anyone who is familiar with working in excel sheets. This will be more beneficial for an average person who is not familiar with advanced technologies.

The Architecture of Data Mining Add in for SQL Server 2008

Following diagram illustrates the architecture of the data mining add in for SQL Server 2008 R2 and its components.[14]

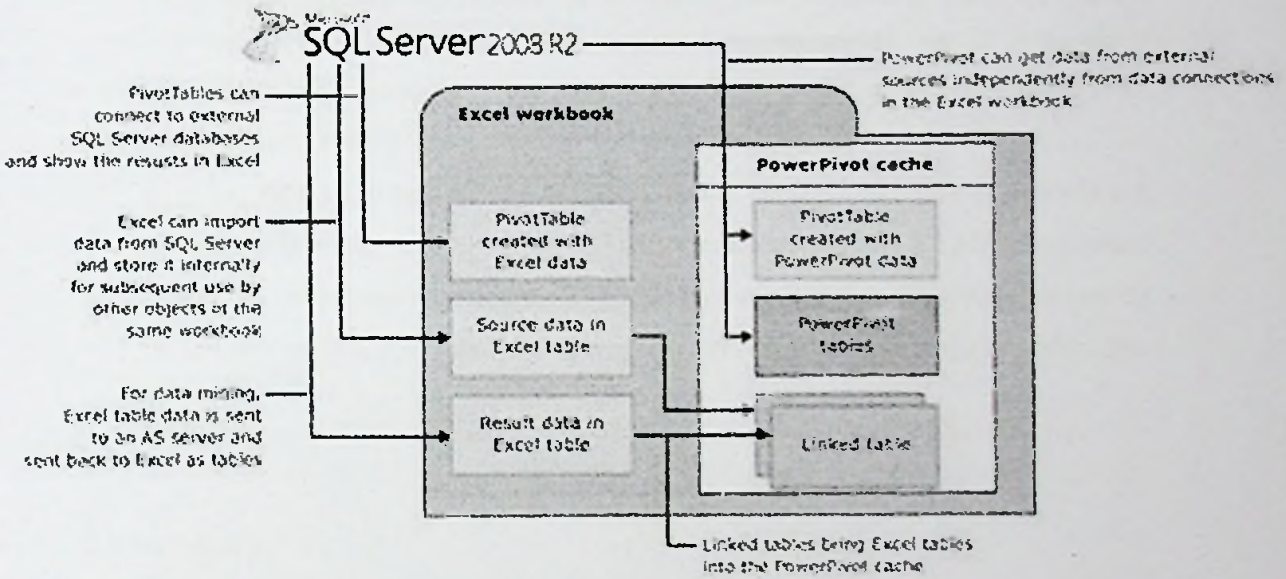

Figure 8: The Architecture of Data Mining Add in for SQL Server 2008

3.6.2 Microsoft .NET Architecture for Web Application Development

For the application development stage Microsoft.NET framework has been used. The reason for selecting this technology is the easy and the awareness of the language. It is having lots of tools and controls for easy development of applications.

The connectivity for the SQL Server database is very reliable and easy to develop inside .NET architecture using inbuilt class libraries. For example to connect to the database there are lots of class libraries like ODBC,OLEDB are provided with lots of database functionalities like connection creations, reading datasets,datareaders,commands ect.

There are lots of functionalities provided for reporting and charting purposes. At the analysis part of the research I have used different kinds of reports for comparing, analyzing and

predicting of electricity consumption patterns of domestic consumers. Basically Microsoft RDLC report component is used for this requirement.

Developing Object Oriented Programming is very easy in .NET framework, where you can have different kinds of Class Libraries and Data Definition Languages to make Object Oriented Concepts in a layered architecture

3.6.3 C# (ASP.NET) Programming Language in Visual Studio 2013

C# for ASP.NET is drastically reduces the amount of code required to build applications within Visual Studio 2013 Integrated development Environment[15]. As C# is an object oriented programming language, it's easy to do the module programming.ASP.NET framework is complemented by a rich tool box and designer in the Visual Studio 2013 environment. Editing, dragging and dropping server controls like powerful tools are provided hence it's easy to do coding[13.The source code and the HTML pages are e write easy to write and maintain. The deployment of the web application to application servers like IIS (Internet Information Server) is easy to manage within few clicks

3.7 Summary

This chapter summarizes the technology adopted for the research and the reasons for choosing that technology in a detailed manner. The next chapter will describe the approach for the research by using the mentioned technologies.