Forecasting Crude Oil Price Using Time Series Data Analysis and Artificial Neural Networks

By

PC Uzma Zafar
NC Muhammad Hanzla Mateen
NC Muhammad Ali Raja
NC Samar Munir

Submitted to Faculty of Computer Software Engineering,

National University of Sciences and Technology, Islamabad in partial fulfillment for

the requirement of a B.E Degree in Computer Software Engineering

JUNE 2014

ABSTRACT

This undergraduate dissertation describes the project which implements a methodology for forecasting crude oil price. This project is a Web-based Application, the user will communicate with the system through web pages using web browser. In this 3-tier architecture Project there is a server-side component which will be responsible for database and synchronization services. The system is based on three main processes as historical data collection, forecasting and results (Graphs). The scope of the project encompasses both server and client-side functionalities.

The web application maintains a data warehouse for storing all the data required by the forecasting tool for the prediction of oil prices. The data is downloaded in the form of Excel files from online sources and is loaded into the Staging Database. Once the data is loaded into the Staging Database, cleansing is performed on the data i.e. data is normalized and rearranged. After the cleansing of data, the data is loaded into the Data-vault Database. This is a non-volatile (permanent) data storage. From here the data is fetched for Neural Network training, Forecasting, Reporting and generating Graphs.

JavaScript and AJAX are used on client side as scripting language. On server side, Windows Server 2012 is used as operating system and database is managed using Microsoft SQL server 2012. Asp.net framework 4.5 is used as server side scripting language.

CERTIFICATE OF CORRECTNESS AND APPROVAL

Certified that work contained in this thesis "**Forecasting Crude Oil Price Using Time Series Data Analysis and Artificial Neural Networks**" carried out by Muhammad Hanzla Mateen, Uzma Zafar, Muhammad Ali Raja and Samar Munir, under the supervision of Dr Seemab Latif for partial fulfillment of Degree of Bachelor of Software Engineering is correct and approved.

APPROVED BY

_____

(Supervisor Name)

_____DEPARMENT

MCS

DATED: _____

DECLARATION

We declare that the work presented herewith is the result of sole effort of our group, comprising of Muhammad Hanzla Mateen, Uzma Zafar, Muhammad Ali Raja and Samar Munir and is free of any kind of plagiarism in part. We also declare that the dissertation has never been submitted previously in part or whole in support of another award or qualification either at this institution or elsewhere.

DEDICATION

In the name of Allah, the Most Merciful, the Most Beneficent

To our parents, without whose unflinching support and cooperation,

a work of this magnitude would not have been possible.

ACKNOWLEDGMENT

We are grateful to Almighty Allah for bestowing us with the strength to undertake and complete the project.

We owe a special debt of gratitude to our supervisor, Dr. Seemab Latif for putting us on right track and guiding us in understanding the concepts of Oil Price trends using latest techniques and for the continuous supervision, motivation and support provided to us right through the project. Without her supervision we would not have been able to complete this successfully. We would like to thank all our other teachers for guiding us in solving our problems related to our project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1. Introduction:

The purpose of this document is to present a detailed description of the requirements, functionalities and testing of our project, ***Forecasting Crude Oil Price Using Time Series Data Analysis and Artificial Neural Networks***. This document will cover each of the system's intended features, what the system will do, the constraints under which it must operate, how the system will behave when operated by user, as well as offer a preliminary glimpse of the software application's User Interface (UI). The document will also cover hardware, software, and various other technical dependencies along implementation and testing details.

1.1 Document Conventions:

The format of this document is simple, divided up into sections detailing different aspects of the system. Bold face and indentation is used on general topics and on specific points of interest. The remainder of the document will be written using the standard font, New Times Roman. Every requirement statement is assumed to have its own priority as to define in most appropriate way the system behavior. In addition there are various figures that represent the described system, where it is needed, and serve only for better understanding of the deployment.

1.2 Background:

Oil plays an increasingly significant role in the world economy since nearly two-thirds of the world's energy consumption comes from crude oil and natural gas. Worldwide consumption of crude oil exceeds $500 billion—roughly 10% of US GDP, and crude oil

is also the world's largest and most actively traded commodity, accounting for about 10% of total world trade.

The crude oil price is basically determined by its supply and demand, and is strongly influenced by many irregular past/present/future events like weather, stock levels, GDP growth, political aspects and people's expectations. Furthermore, since it takes considerable time to ship crude oil from one country to another, oil prices vary in different parts of the world.

In addition, as sharp oil price movements are likely to disturb aggregate economic activity, volatile oil prices are of considerable interest to many researchers and institutions. Therefore, forecasting oil prices is a very important topic, although it is an extremely hard one due to its intrinsic difficulty and practical applications.

Much of the work on forecasting the price of oil has focused on the dollar price of oil. This is natural because crude oil is typically traded in U.S. dollars.

1.3 Problem Domain:

The concept of this software originates from the idea of giving the user such an environment in which the information gained from the system can be used by the user in making healthier business decisions. The user will be able to query the desired information from the system, and that query will be answered in the form of information.

For implementing such a software, proper data should be collected and stored in a data warehouse and appropriate forecasting method should be applied on this data to generate results.

Thus, three important concepts are involved:

i. Data Warehousing
ii. Forecasting
iii. Data Mining

1.4 Project Vision:

For	People, Organizations affected by Oil Prices
What	To forecast Oil Prices and perform reporting on historical data.
The	Oil Price Forecasting System
Is	Web Application
That	Provides efficient, convenient reporting using graphical representation along with daily oil price forecasts.
Unlike	Monthly oil price forecasts
Our Product	Ensures accurate price forecasts and provides convenient querying of data using graphs and charts.

Table 1- 1 Project Vision

1.5 Goals and Objectives:

Following are the academic objectives and goals of the project achieved.

1.5.1 Academic Objectives:

i. To learn the concepts of Forecasting, Time Series Analysis from Online resources.
ii. Pattern Extraction.
iii. To practice all aspects of Software Engineering i.e. Software Requirement Engineering, Software Design and Architecture, Software Project Management, Software Quality Assurance and Testing.

1.5.2 End Goal:

i. Collect Data
ii. Create a Data Warehouse, which stores, all the relevant information regarding the Oil prices.
iii. Extract Transform and Load, a technique that is going to shift all the different formats of data into one single format of the data.

iv. Use appropriate Model for Forecasting.
v. Perform Time Series Analysis/ Machine Learning Analysis.
vi. The End Goal of this Project is to develop Web Application, which forecasts the price of Crude Oil using Time Series Data Analysis.

1.6 Organization

The first part of thesis is the abstract which describes the main details of the project, followed by the introduction section which specifies the problem statement, approach, scope and objectives. The literature review section state the various resources read online before the commencement of the project. The design and development part illustrate the diagrams which describe the detailed design of the system, its components, interfaces and data necessary for the implementation phase. The analysis and evaluation part give details of the black box testing, unit testing and system integration testing; actual results against expected results.

1.7 Deliverables:

Following are the application and document level deliverables prepared during project.

1.7.1 Document Deliverables:

i. Software Requirement Specification.
ii. Architecture and Design Document.
iii. Project documentation and Code.
iv. Testing Document.
v. User Manual.

1.7.2 Application Level Deliverables:

i. Data about the variables, which are considered for the forecasting of crude oil price, in .xsls format.
ii. Data warehouse creation and management using Extract Transform Load Tool.

iii. Applying forecasting technique on historical data.

iv. Website designing and adding functionality for oil price prediction of defined algorithm.

v. Generate graphs and charts as output according to result of algorithm.

CHAPTER 2
LITERATURE REVIEW

2. Literature Review:

2.1 Data Warehousing:

Data warehouse is a database used for reporting and data analysis, which store current as well as historical data and are used for creating trending reports.

2.1.1 Basic Processes of the Data Warehouse:

2.1.1.1 Extracting

Reading and understanding the source data, and copying the parts that are needed to the data staging area.

2.1.1.2 Transforming

i. Cleaning data (correcting, resolving conflicts, dealing with missing data, etc.)
ii. Purging data (eliminating extracted data not useful for data warehousing)
iii. Combining data sources (matching key values, fuzzy matches on non-key values, etc.)
iv. Restructuring the data (so it confirms to the structure of the target DWH)
v. Creating surrogate keys (in order to avoid dependence on legacy keys)
vi. Building aggregates

2.1.1.3 Loading

i. Bulk loading

2.2 Forecasting:

Forecasting is the process of making statements about events whose actual outcomes (typically) have not yet been observed. A commonplace example might be estimation of some variable of interest at some specified future date.

The basic steps in a Forecasting Process are:

i. Identify the Goal of the Forecast
ii. Establish a Time Horizon
iii. Collect Data
iv. Select a Forecasting Technique
v. Conduct the Forecast
vi. Monitor Accuracy

2.3 Data Mining:

Data Mining is an analytic process designed to explore data in search of consistent patterns and then to validate the findings by applying the detected patterns to new subsets of data. The ultimate goal of data mining is prediction.

2.3.1 Methodology:

The process of data mining consists of three stages:

2.3.1.1 Exploration:

Identify the most relevant variables and determine the complexity and/or the general nature of models that can be taken into account in the next stage.

2.3.1.2 Model building and validation:

This stage involves considering various models and choosing the best one based on their predictive performance. There are a variety of techniques developed to achieve that goal - many of which are based on so-called “competitive evaluation of models”.

2.3.1.3 Deployment:

That final stage involves using the model selected as best in the previous stage and applying it to new data in order to generate predictions or estimates of the expected outcome.

2.4 Previous and Latest Work done in Problem Domain:

Analysis shows that there are a lot of methods and techniques for Oil Price Predictions. A number of different methods have been applied in order to predict the Oil Prices.

2.4.1 Categories of Forecasting Methods:

There are many methods which are being used for predicting Oil Prices and Trends. These methods can be grouped in these major categories:

i. Naïve Approach
ii. Time Series Methods
iii. Econometric Forecasting Methods
iv. Judgmental Methods
v. Artificial Intelligence Methods
    i. Data mining
    ii. Machine Learning
    iii. Pattern Recognition

2.5 Papers and Reviews/ Surveys:

The various papers and articles read, gave the information about various methods and pattern learning techniques applicable for prediction Oil Prices and Trends. Following are summaries of few research papers and techniques which are relevant to the problem domain.

2.5.1 Prediction of Oil Production with: Data Mining, Neuro-Fuzzy and Linear Regression

In this paper the main idea of the author is to do oil forecasting by using Data Mining Technique for improving estimation of oil consumption base of history of data. The implementation of auto regression, Data cleaning and concept of pre-processing to viewpoint of Time series analysis has been used.

The paper concludes that ANFIS algorithms can be considered in terms of simplicity, generality and applicable. In this paper an accurate method to forecast oil production is presented. The proposed efficient, simple and fast method works based on ANFIS algorithm that cleans and integrates data and then turns to static all data, then divide data to two categories (train data and test data) after this create different ANFIS network and train networks with train data. Finally, evaluating test data to all train networks and choose the network that has lower error among all networks, and call it Best-network.

2.5.2 Crude Oil Price Forecasting with TEI@I METHODOLOGY*

Wang and Yu proposed a new methodology for handling complex systems that is **TEI@I methodology** by means of a systematic integration of text mining, econometrics and intelligent techniques. The crude oil price data used in this study are monthly spot prices of West Texas Intermediate (WTI) crude oil, covering the period from January 1970 to December 2003 with a total of n= 408 observations.

The simulation results show that the proposed nonlinear integrated forecasting approach with error correction and judgmental adjustment produces a substantial improvement in oil price forecasting and the nonlinear integrated forecasts are superior to the simple integrated forecasts and the individual forecasts alone. These two conclusions lead to a final conclusion: the novel nonlinear integrated forecasting model can be used as an alternative tool for crude oil price forecasting to obtain better forecasting accuracy.

2.5.3 Bayesian Neural Networks for Nonlinear Time Series Forecasting

In this paper, author apply BNNs to time series analysis, and propose a Monte Carlo algorithm for BNN training. In addition, he go a step further in BNN model selection by putting a prior on network connections instead of hidden units as done by other methods.

In this article, we only want to show that BNN outperforms many competitors in forecasting nonlinear time series, although article also provide a Monte Carlo algorithm for BNN training. Theoretically, any Monte Carlo method, if it can sample from the posterior effectively, will have the same forecasting performance as the evolutionary Monte Carlo (EMC) algorithm provided in this article.

2.5.4 Econometric Models for Oil Price Forecasting

According to this article, for oil price forecasting one can distinguish among three categories of econometric models:

i. Time series models exploiting the statistical properties of the data, namely autocorrelation and non-stationary.
ii. Financial models based on the relationship between spot and future prices; and
iii. Structural models describing how specific economic factors and the behavior of economic agents affect the future values of oil prices.

This article was helpful in identifying the variables that affect oil prices. Besides the impact of OPEC, many authors have also recognized the importance of the current and future availability of physical oil. According to this view, the most crucial variable is represented by the level of inventories. Stocks are the link between oil demand and production and, consequently, they are a good measure of price variation.

CHAPTER 3
SOFTWARE REQUIREMENT SPECIFICATION

---

3. Software Requirement Specification:

3.1 Project Scope:

The aim of our project is to design an application which Forecasts the Oil Prices using Time Series Analysis for multiple users. One of the main advantages that our product will provide is the effective usage of "Dashboards" through which the user will be able to identify trends, as the graphical representation of any system helps best in understanding the problem. Another key aspect of our system is that it gives users a broad control over what kind of information they want, such as day to day information will also be provided to the users.

Modules that come under project scope are:

i. **Historical Data (Data Collection):** It includes extracting the required information and analyze data from bulk of raw data and creating Data warehouse using this data.

   The variables that are being considered are:

   i. CPI
   ii. Oil Price
   iii. Oil Production
   iv. Oil Supply
   v. Oil Stock

Also, factors which are being considered as constants are:

   i. Value of the Economical Factor
   ii. Value of Law and Order (Security)
   iii. Value of the Political Factor

Once the data has been collected, Data warehouse is created which stores all the historical data.

ii. **Prediction Algorithm (Analysis):** It includes predicting Oil Price trends using Artificial Neural Networks and Time Series Data Mining Algorithm.

iii. **Results of Prediction (Presentation):** It includes generating Graphs and Report of analysis of prediction results.

3.2 Design and Implementation Constraints:

Constraint	Description
Economic	None
Political	The system will abide by the rules and regulations of the Organization/Institution.
Technical	Only C# can be utilized to create this software, due to requirement of web form based design. Asp. Net framework, Ajax, JavaScript. So the browsers shall support Ajax and JavaScript for full functionality. SQL database will be required at server end. Internet Connection is required.
Security	No Security Constraints.
Schedule & Resources	Project needs to be completed in under eight months.

Table 1- 2 Design and Implementation Constraints

3.3 User Documentation:

i. The system will provide an online user manual in HTML that describes the functionality and options available to the user.

ii. The system will provide a hard copy of the user manual, which is identical to the Online manual.

3.4 Assumptions and Dependencies:

3.4.1 User related:

i. The users have sufficient knowledge of Computers.
ii. The users know the English language, as the user interface will be provided in English.

3.4.2 Bandwidth related:

The computer shall have efficient internet facility and browsing capabilities.

3.4.3 Time Dependencies: (None Required)

3.4.4 Hardware/Software Dependencies: (None Required)

3.5 System Features

The following list offers a brief outline and description of the main features and functionalities of our system.

i. **Web Application:** Application is developed using ASP.net and C#, it provides an interface in the application.

ii. **User Query:** The user enters the desired query and then the results are displayed.

iii. **Data Warehouse:** A comprehensive database is maintained on the Server side which has all the necessary data about previous Oil prices.

iv. **Price Forecasting:** Oil Prices are forecasted applying Time Series Analysis on previous stored data.

v. **Dashboards:** Dashboards are used to give an overview of how properties are being performed by displaying summaries of different reports as widgets on a single web page.

vi. **Graphical User Interface**: The Interface of the application has different views for displaying the result.

vii. **User authentication:** Allows different users to login to the system.

viii. **Manage User accounts**: This includes operations such as add, modify and delete users from the database.

ix. **Maintains Logs:** It maintains all the history of previous forecasts and other usage.

x. **Report Generation:** The system generates the output in the form reports, when requested.

3.5.1 Forecast Oil Price Value:

3.5.1.1 Description and Priority:

The user requests the System to generate/predict Oil price.

Priority: High.

3.5.1.2 Stimulus/Response Sequences:

Stimulus: The user wants to forecast the Oil Price.

Response: The System displays the forecasted price.

3.5.1.3 Functional Requirements:

RE-1: Online Server: The server must be at all times running in order to get the results, as all calculations as well as data gathering will be done at the server end.

RE-2: Database Connected to Server: A connection must be established at the sever end between server and database so server gets desired values to process upon.

RE-3: Browser Compatibility: The browser being used by the user must be compatible to render ASP.Net page(s).

3.5.2 Maintain Database with newly Predicted Value(s):

3.5.2.1 Description and Priority:

When the user has predicted an outcome, that outcome will also be stored in the database for future use.

Priority: Low

3.5.2.2 Stimulus/Response Sequences:

Stimulus: The user wants to forecast the Oil Price.

Response: The System queries the user.

Stimulus: System displays predicted outcome.

Response: System records the value in the database.

3.5.2.3 Functional Requirements:

RE-1: Connection of the Database and the product.

3.5.3 User Login:

3.5.3.1 Description and Priority:

Administrator and the other members of the organization require to login to the system on each first access. After login they can do the further actions accordingly.

Other users who want to use the purposed System need to first register in the System through the admin and after this he/she can use the Web Application Features.

Priority: High

3.5.3.2 Stimulus/Response Sequences:

Stimulus: The user accesses the login page.

Response: The System asks the user to enter his username and password.

Stimulus: The user enters his username and password and presses submit.

Response: The System verifies the credentials from the server and either allows the user access or provides an error message of invalid details.

Stimulus: The user acce0sses the login page and logs in as guest.

Response: The System allows the user to view graphs.

Stimulus: If the user does not have an account.

Response: The System allows the user to create an account and asks him for following information:

i. Name of the User
ii. Email of the User
iii. Username
iv. Password

3.5.3.3 Functional Requirements:

RE-1: The System will check the validity of the user from database (which has the passwords and usernames already stored).

RE-2: The System should provide an interface for the user to login.

RE-3: Online Server: The server must be at all times running in order to get the results, as all calculations as well as data gathering will be done at the server end.

RE-4: Database Connected to Server: A connection must be established at the sever end between server and database so server gets desired values to process upon.

RE-5: Browser Compatibility: The browser being used by the user must be compatible to render ASP.net pages.

3.5.4 View Graphs of trends in Oil Price:

3.5.4.1 Description and Priority:

The user requests the system to generate graph(s) for viewing Oil price trends of a particular year or range.

Priority: High

3.5.4.2 Stimulus/Response Sequences:

Stimulus: The user enters only the year and clicks on the generate graph button.

Response: The system prompts the user on the type of graph the user wants to display.

Stimulus: The user presses the generate button, without entering any specific year/range.

Response: The system prompts an error message.

3.5.4.3 Functional Requirements:

RE-1: Online Server: The server must be at all times running in order to get the results, as all calculations as well as data gathering will be done at the server end.

RE-2: Database Connected to Server: A connection must be established at the sever end between server and database so server gets desired values to process upon.

RE-3: Browser Compatibility: The browser being used by the user must be compatible to render ASP.Net page(s).

3.5.5 Administrator Role on Server Side:

3.5.5.1 Description and Priority:

Administrator has the right to manage the databases and accounts of users.

Priority: High

3.5.5.2 Stimulus/Response Sequences:

Stimulus: The admin accesses the login page on the server side.

Response: The System asks the admin to enter Login details.

Stimulus: The admin enters his username and password and presses submit.

Response: The System verifies the credentials from the Server and either allows the admin access or provides an error message of invalid details.

Stimulus: The admin is logged in successfully.

Response: The System allows admin to:

i. Manage Account
    i. Create Account.
    ii. Update Account.
    iii. Delete Account.
ii. Manage Database
    i. Update Database.
    ii. Delete Database.
    iii. Create Database.

Stimulus: The admin selects Account Creation.

Response: The System asks him for following information:

i. Name of the User
ii. Email of the User
iii. Password

Stimulus: The admin selects Update Database.

Response: The System allows admin to update the following databases:

i. Staging Database.
ii. DataVault Database.
iii. User Database.

Stimulus: The admin selects Delete Database.

Response: The System allows admin to delete the following databases:

i. Staging Database.
ii. DataVault Database.
iii. User Database.

3.5.5.3 Functional Requirements:

RE-1: The System will check the validity of the admin from database (which has the passwords and usernames already stored).

RE-2: The System should provide an interface for the admin to login.

3.5.6 Report Generation:

3.5.6.1 Description and Priority:

The user requests the System to generate a report.

Priority: High.

3.5.6.2 Stimulus/Response Sequences:

Stimulus: The user wants to generate a report.

Response: The System queries the user to enter the details about the report.

Stimulus: The user enters the details and presses the generate button.

Response: The System shows the desired result on the next page.

Stimulus: The user presses the generate button, without entering any specific date/detail.

Response: The System prompts an error message.

3.5.6.3 Functional Requirements:

RE-1: Online Server: The server must be at all times running in order to get the results, as all calculations as well as data gathering will be done at the server end.

RE-2: Database Connected to Server: A connection must be established at the sever end between server and database so server gets desired values to process upon.

RE-3: Browser Compatibility: The browser being used by the user must be compatible to render ASP.Net page(s).

3.5.7 Manage Profile:

3.5.7.1 Description and Priority:

The user can edit the account details and view log.

Priority: Low.

3.5.7.2 Stimulus/Response Sequences:

Stimulus: The user is logged in successfully and wants to edit the account details.

Response: The System allows user to:

i. View Log.
ii. Manage Account.
    i. Update Account.
    ii. Delete Account.

Stimulus: The user selects Update Account.

Response: The System allows the user to edit the following information:

i. Username
ii. Email
iii. Password

Stimulus: The user selects Delete Account.

Response: The System deletes the account.

Stimulus: The user selects to view log.

Response: The System displays log.

3.5.7.3 Functional Requirements:

RE-1: Online Server: The server must be at all times running in order to get the results, as all calculations as well as data gathering will be done at the server end.

RE-2: Database Connected to Server: A connection must be established at the sever end between server and database so server gets desired values to process upon.

RE-3: Browser Compatibility: The browser being used by the user must be compatible to render ASP.Net page(s).

3.6 Nonfunctional Requirements

3.6.1 Performance Requirements:

3.6.1.1 Capacity:

i. The system shall be capable of handling minimum of 20 users at time.
ii. **Mean Time to Repair (MTTR):** Even if the system fails, the system shall be recovered back up within 30 minutes.

3.6.1.2 Throughput:

The response time for an average transaction tends to decrease with the increase in overall throughput. The server will be able to handle 20 simultaneous requests.

3.6.2 Safety Requirements:

Database backup will be created in order to keep the data record safe and updated in case of any database crash.

3.6.3 Security Requirements:

The web app will be a solely read only web app. No changes will be able to be made from the user end.

3.6.4 Software Quality Attributes:

3.6.4.1 Availability:

The system shall be available to the user 9 out 10 times and pop up error messages in case of component failure. In that case the error messages appear when something goes wrong so to prevail availability problems.

3.6.4.2 Maintainability:

The maintenance of the system shall be done after every 6 months such that there is no unnecessary data stored on server.

3.6.4.3 Design Quality:

3.6.4.3.1 Evolution:

The Application shall have the capability to evolve according to the changing needs/environment. It shall be able to incorporate new requirements. Therefore, system shall be checked for evolutionary changes every year.

3.6.4.4 Usability:

Checking that the system shall be easy to handle and navigates in the most expected way with no delays. In that case the system program reacts accordingly and transverses quickly between its states.

3.6.4.4.1 Learnability:

The system is fairly easy to use with clearly labeling and an info button on each page. A new user will require at maximum 1 day understanding how the entire system is to be operated. The system will provide an online user manual that describes the functionality and options available to the user.

3.6.4.4.2 Efficiency:

The Webpage shall be displayed in no more than 10 seconds using a 500 - 800 Kbps Bandwidth connection per Machine/Computer.

3.6.4.4.3 Flexibility:

The design and architecture of the application shall be flexible enough for catering any new requirements, if any at some later stage or for the application enhancement.

3.6.4.5 Functionality:

Checking that the system provide the right tools for processing the user query and sending it to server side, in short checking whether the system performs its intended functions.

3.6.4.6 Reliability:

This system should be reliable and provide accurate data without errors. Since it is a web based system, the system is highly portable and can be accessed from anywhere with an internet connection.

3.6.4.7 Robustness

After the failure occurs, the system will be able to recover within 2 hours after restart the system.

CHAPTER 4
SOFTWARE DESIGN SPECIFICATIONS

4. Architecture

4.1 System Design:

Following is the architecture design of the system.

4.1.1 Architecture Approach:

This system is a Web Application, which is implemented using three layer architecture pattern with traditional asp.net web form methodology.

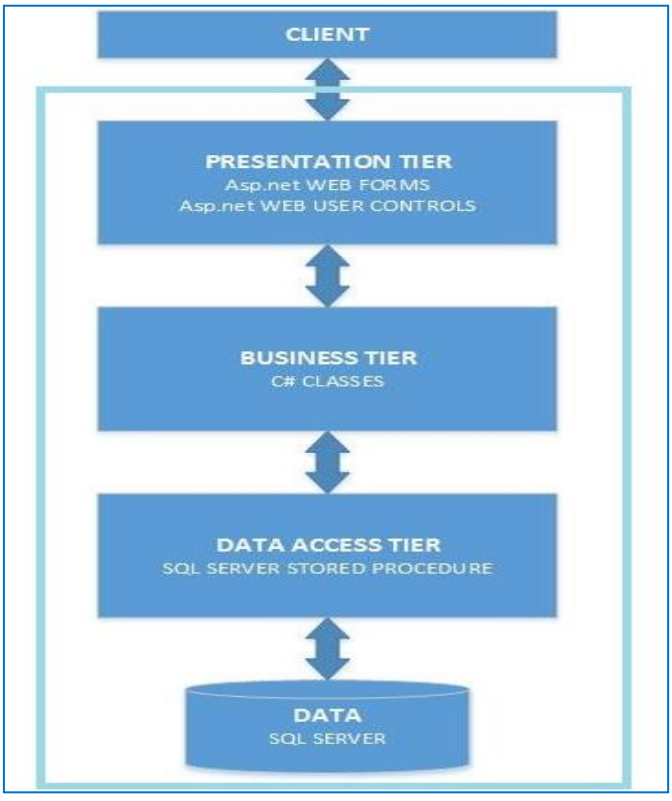

Figure 4-1 3-tier Architecture Diagram

4.1.1.1 Description:

i. **Presentation Layer:** Presents the results to the users. It interacts with business logic layer to fetch data and pass user data for processing.

ii. **Business Logic Layer:** This layer performs the business logic processing that occurs. All options in presentation layer will interact with relevant modules in this layer which further will interact with data access layer for data handling. These are actually the classes dealing with logic in functions.

iii. **Data Access Layer:** Data Access Layer handles the database processing and the access to the data. In the figure above, the Data Context module deals with retrieving and adding all the data related to application from/in to database server. This has all the classes relevant to database tables.

4.1.2 System Architecture Diagram:

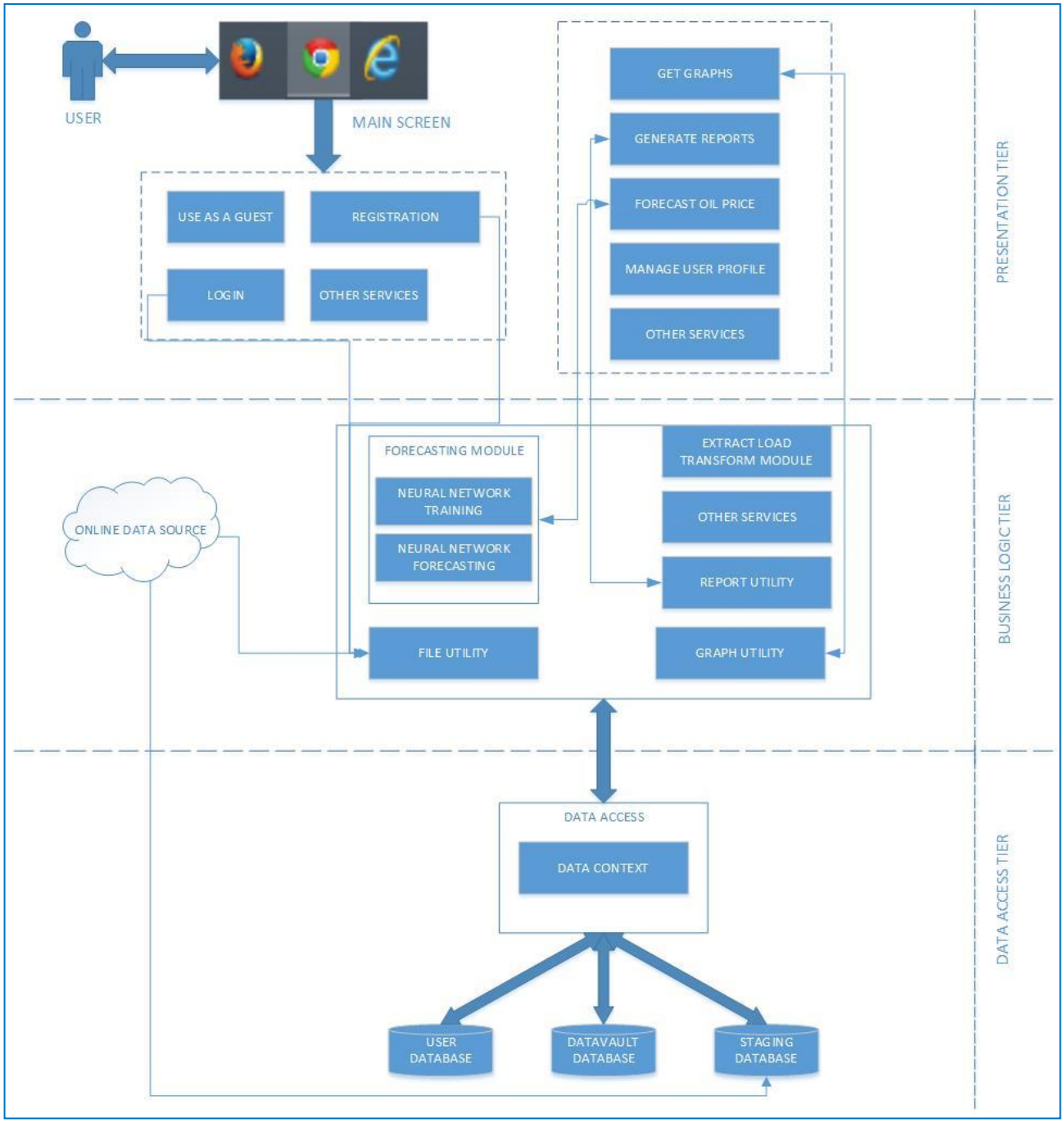

Figure 4-2 System Architecture Diagram

4.1.3 System Block Diagram:

Following block diagram below shows the principal parts of the system and their interactions

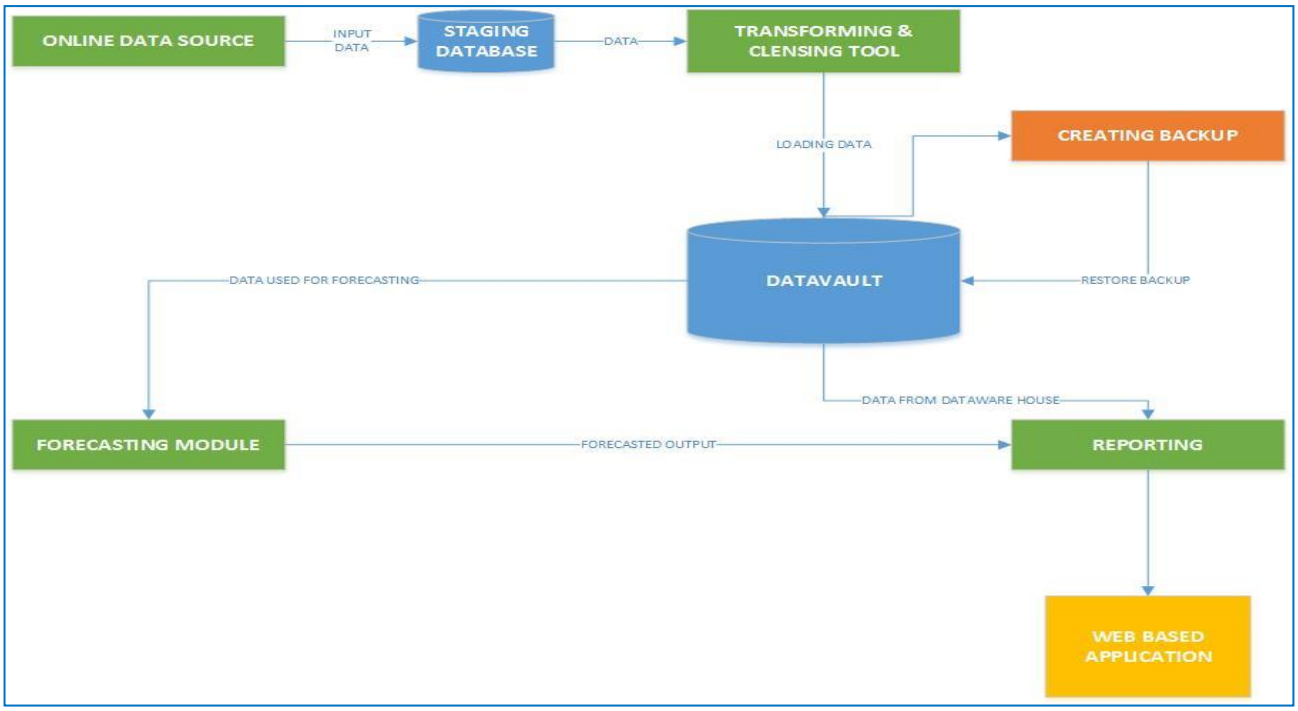

Figure 4-3 System Block Diagram

4.1.4 Perspective Diagram:

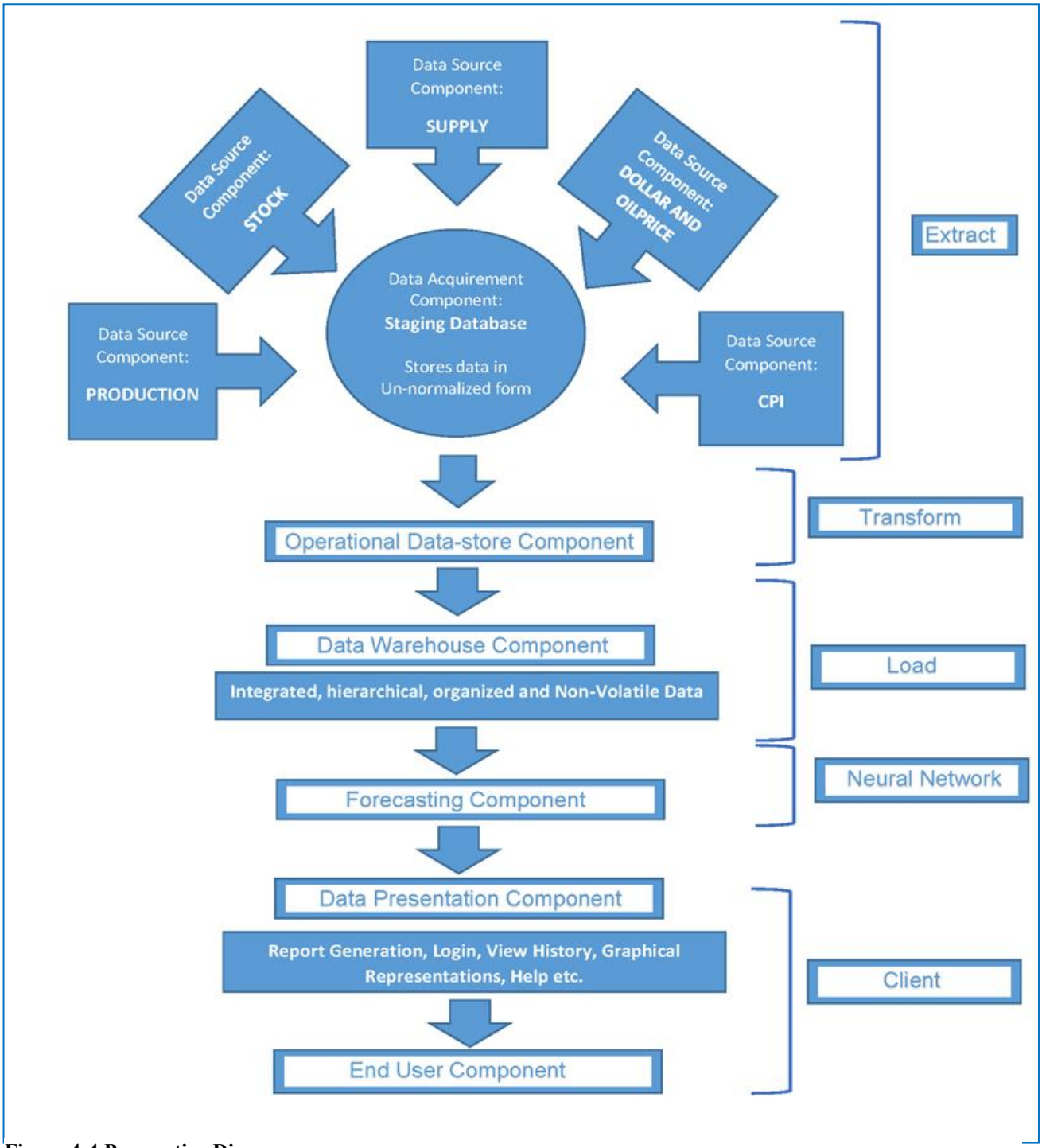

Figure 4-4 Perspective Diagram

4.1.5 Work Breakdown Structure:

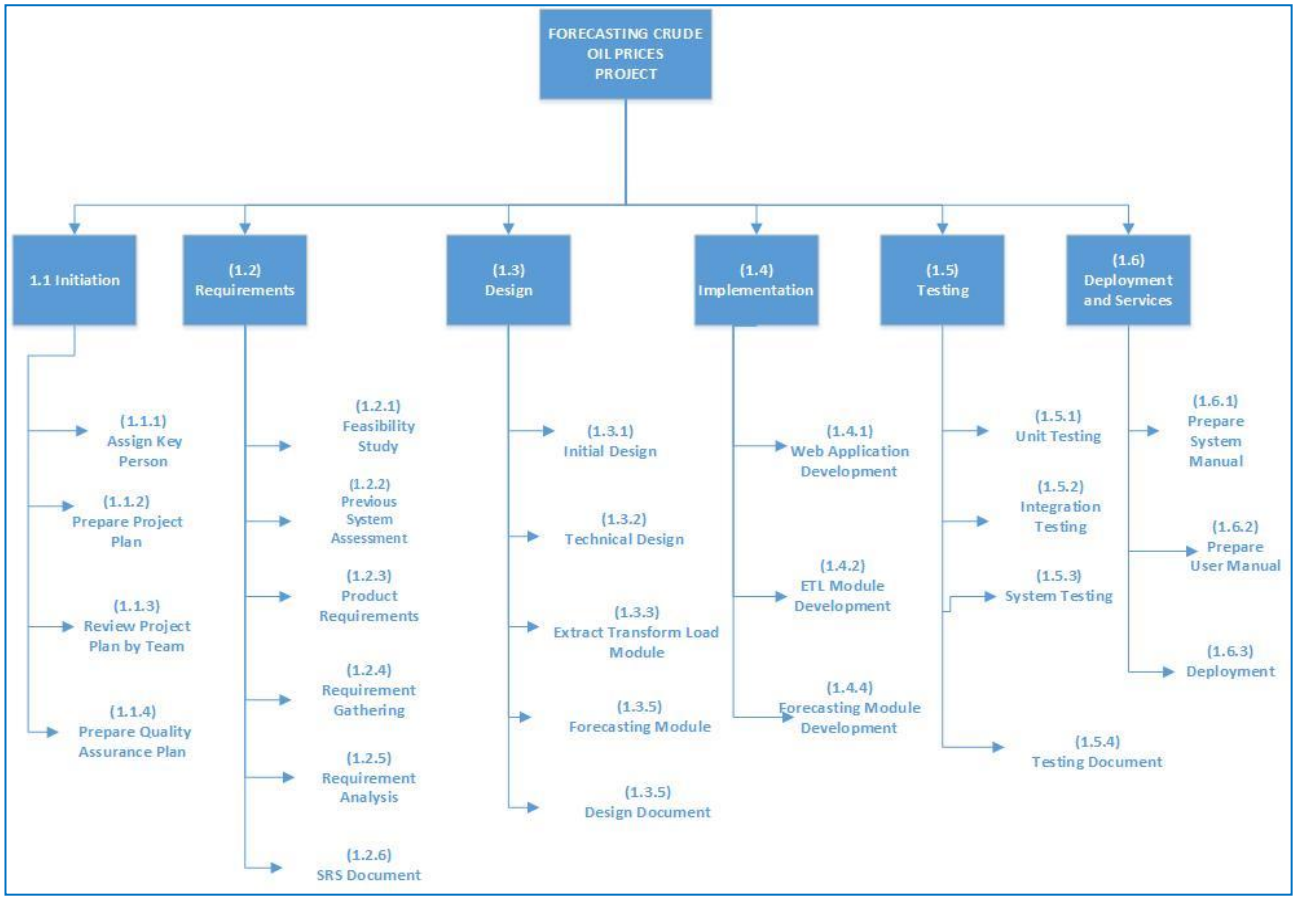

Figure 4-5 Work Breakdown Structure

4.1.6 Brief System Description and Flow:

Our project is a web application. The user is provided with the home page, when he opens the application. The user has two options to access the application, either he can choose to login as a guest or he can provide valid combination of username and password. This pair of username and password is sent to the database for validation. If the combination is correct then the user is successfully logged into the system, and now he can access all the features of the application. If the user logs in as a guest then he can only view graphs, he is not allowed to use the prediction and reporting feature. If the user is new to the system he is allowed to create his account. The user selects the "registration button" and then is allowed to fill in the registration form, when this is done, the corresponding information is sent to database and stored, now the user has username and password to login into the system. Other than this, administrator can login into the server side of the system and is allowed to maintain databases and accounts.

4.1.6.1 When User logs in on Client Side:

The options that a user can select from according to his requirement are:

i. **Get Prediction:**

   User can see the prediction of Oil Prices. The user selects the Get Prediction feature, the system asks the user to enter the date for forecasting the oil price. Also the user selects the values of the constant factors. The user presses the Prediction button. The data is sent to the forecasting tool on the server side. The forecasting tool fetches data from data-vault database. The forecasting tool then forecasts the Oil Price on the specified date and sends this value to the client side and the value is displayed to the user.

ii. **Generate Report:**

   User can generate the reports. The user selects the Get Report feature, the system asks the user to enter the date from/to for generating the Report. The user presses the Report button. The data is sent to the server side, and the data is fetched from data-vault database according to the details specified by the user. Then this data is sent to the client side and is displayed to the user.

iii. Generate Graphs:

User can generate the graphs. The user selects the Get Report feature, the system asks the user to enter the type of graph for generating the Graph. The user presses the Graph button. The data is sent to the server side, and the data is fetched from data-vault database according to the details specified by the user. Then this data is sent to the client side and is displayed to the user, in the form of the graph.

iv. Other Services:

i. Contact Us

ii. Online User Manual

4.1.6.2 When Administrator logs in on Server Side:

Application control panel is managed by administrator. Administrator tasks are to manage the data fetching from online sources, databases, user accounts etc.

4.2 System Decomposition:

The System is decomposed into two main modules, server and client side.

4.2.1 Description:

The major components of the system are the Client Side and the Server Side.

i. The Client is a web browser. The user communicates with the active Server using web application and its purpose is to take input from users, process the input, request the server accordingly and show the results to the users. There can be multiple clients that can connect to a server.

ii. The Server has an ETL (Extract Transport Load) Tool which will be used to create and update the database, which has all the required data. Moreover Financial Predictor Tool will be using Artificial Neural Networks and Time Series Analysis to predict the Oil Prices, using stored data.

4.2.2 User Classes and Characteristics:

For conventional reason we name each of the user classes-actors with this format:

4.2.2.1 Physical Actors:

i. **Administrator:** Administrator is responsible to setup and configure the system before each use. He shall also be able to view and edit records of all users from the database.

ii. **User:** The User is the one that uses the System to predict Oil Prices. The user is allowed to enter the date, and the value of constant factors.

4.2.2.2 System Actors:

i. **Client:** The client is the system that connects to the server and handles the user requests on the session and finally submits the information back to the server.

ii. **Server:** The server is the system that accepts multiple connections from clients and generates the results, which are then shown to users on client side.

iii. **IIS**.

iv. **Database:** It maintains all the required data for the Web Application.

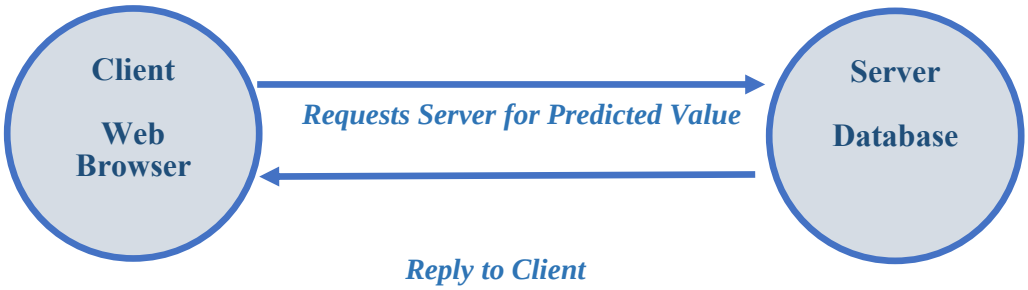

Figure 4-6 Client Server Communication Diagram

4.3 Class Diagram:

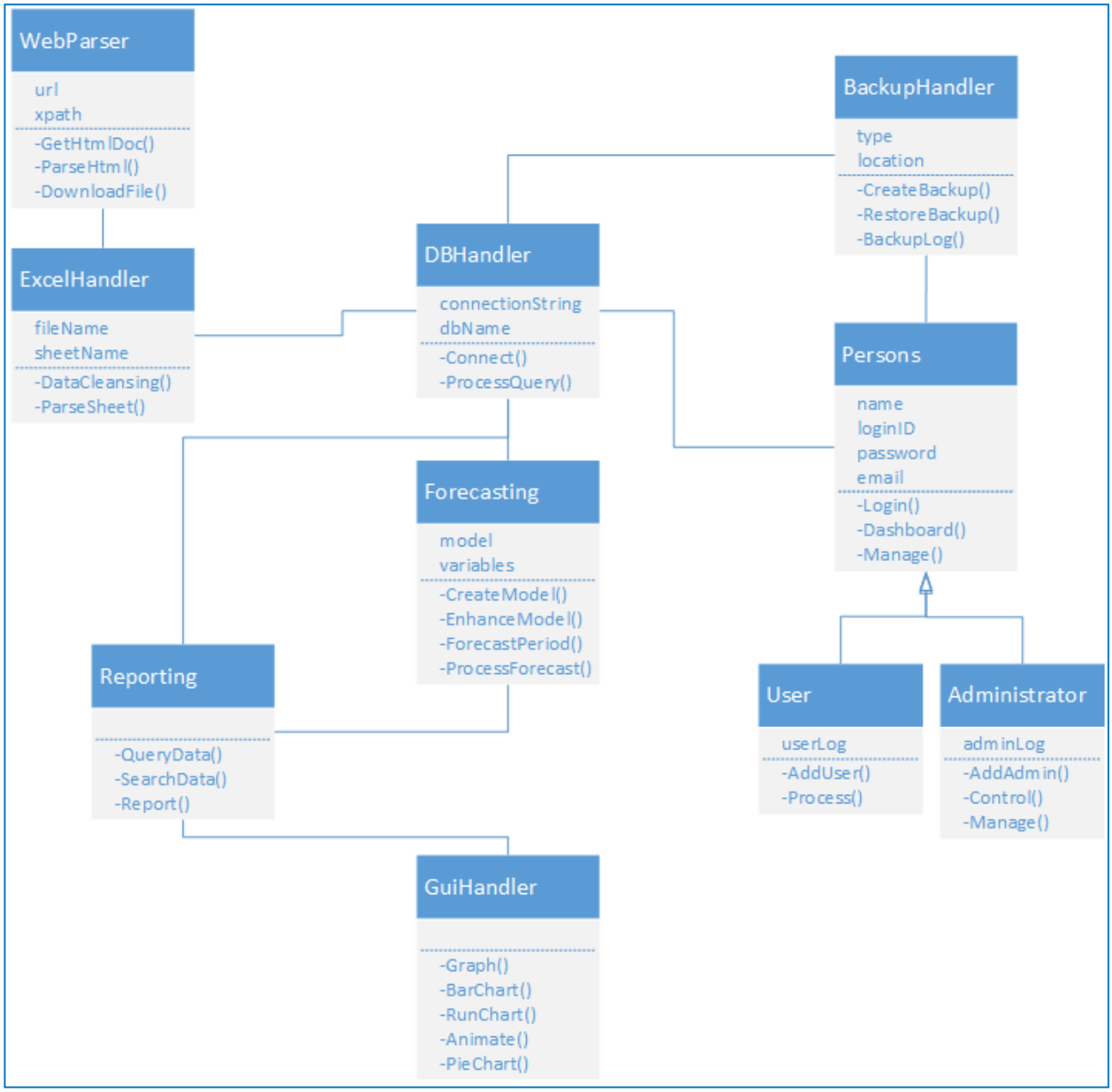

Figure 4-7 Class Diagram

4.4 Server Side:

Server Side has following two main modules:

4.4.1 Extract Transform Load (ETL) Tool:

The web application maintains a data warehouse for storing all the data required by the forecasting tool for the prediction of oil prices. This data warehouse needs to be loaded regularly so that it can serve the purpose of facilitating data analysis. For this purpose, data from various sources is extracted and copied in to the Staging Database of the data warehouse. The main challenge is to integrate, rearrange and consolidate large amount of data, thereby providing a unified information base for business intelligence.

The process of extracting data from sources and bringing it into data warehouse is commonly known as ETL.

i. **Extraction of Data:**

   During extraction, the desired data is identified and extracted from many different online sources, including applications. The data is downloaded in the form of Excel files and is loaded into the Staging Database.

ii. **Transformation of Data:**

   Once the data is loaded into the Staging Database, cleansing is performed on the data i.e. data is normalized and rearranged. Staging database is a volatile database, once the data is loaded onto the data-vault, the data of staging database is deleted.

iii. **Loading of Data:**

   After the cleansing of data, the data is loaded into the Data-vault Database. This is a non-volatile (permanent) data storage. From here the data is fetched for Neural Network training, Forecasting, Reporting and generating Graphs.

4.4.2 Forecasting Tool:

The Forecasting Tool is implemented using the concepts of Artificial Neural Networks. It uses 80% of the data of data warehouse for the training of the network, which uses Back Propagation Algorithm. 20% of the data is used for forecasting the oil prices.

4.4.3 Server Side Flowchart:

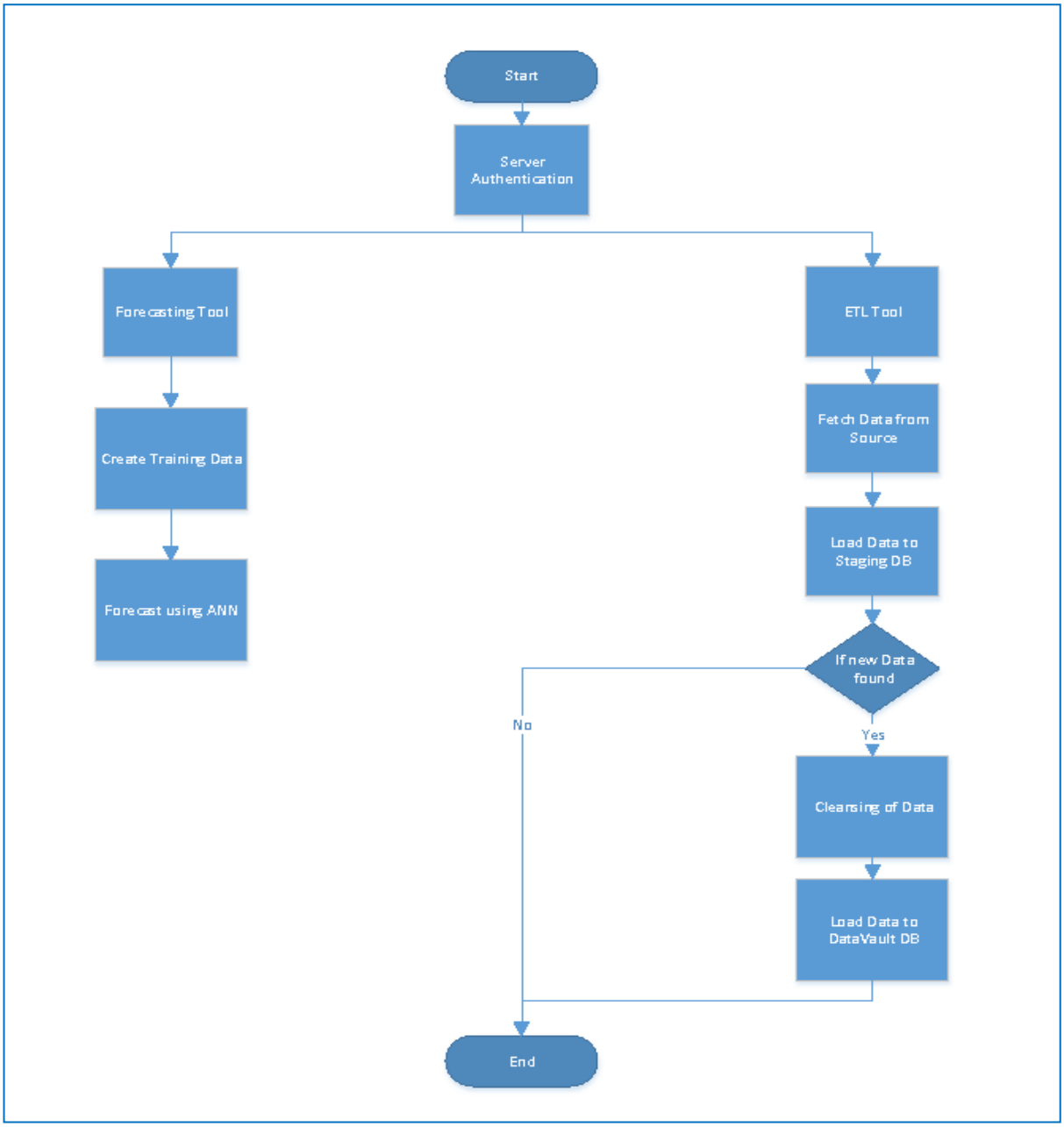

Figure 4-8 Server Side Flowchart

4.4.4 Server Side Use Case Diagram:

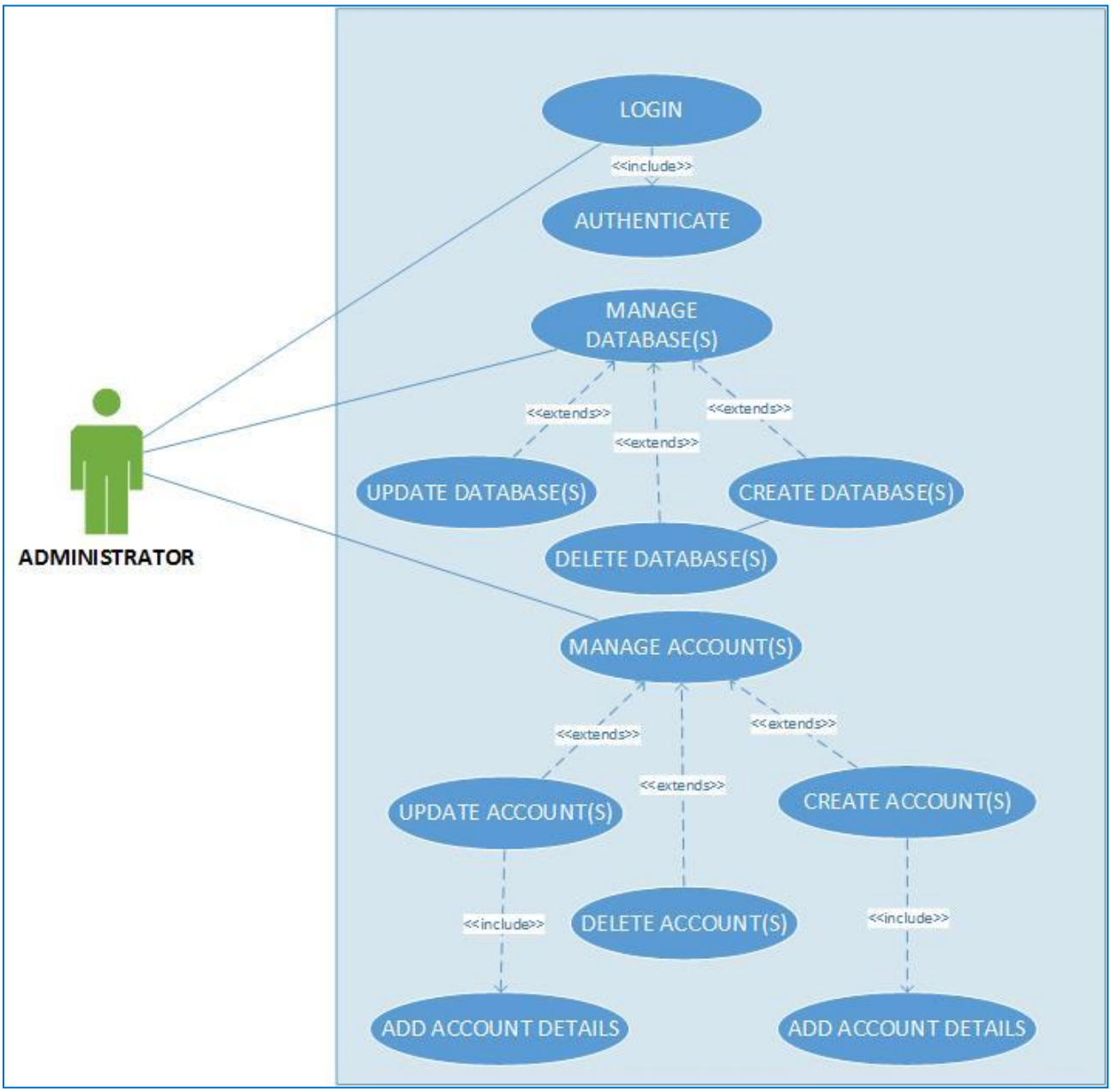

Figure 4-9 Server Side Use Case Diagram

4.5 Client Side:

4.5.1 Software Interface Design:

The interfaces will consist of numerous pages. There will be text fields in which the user can enter the day month and year of the record they want to retrieve may it be a past value or future value. The user can then proceed to generate reports, graphs or report any errors they may be encountered. Each interface will consist of a basic layout which will always have the same background, home button, exit button as well as the button which will take user to the previous page.

User Interface (UI) is designed according to UI design principles.

i. **The structure principle:** UI is organized in such a way that related things are combined together and unrelated things are separated.

ii. **The simplicity principle:** It is easy to follow the provided interface. In the case of mistake, system displays error message.

iii. **The visibility principle:** All system does not overwhelm users with too many alternatives.

4.5.2 Web Pages:

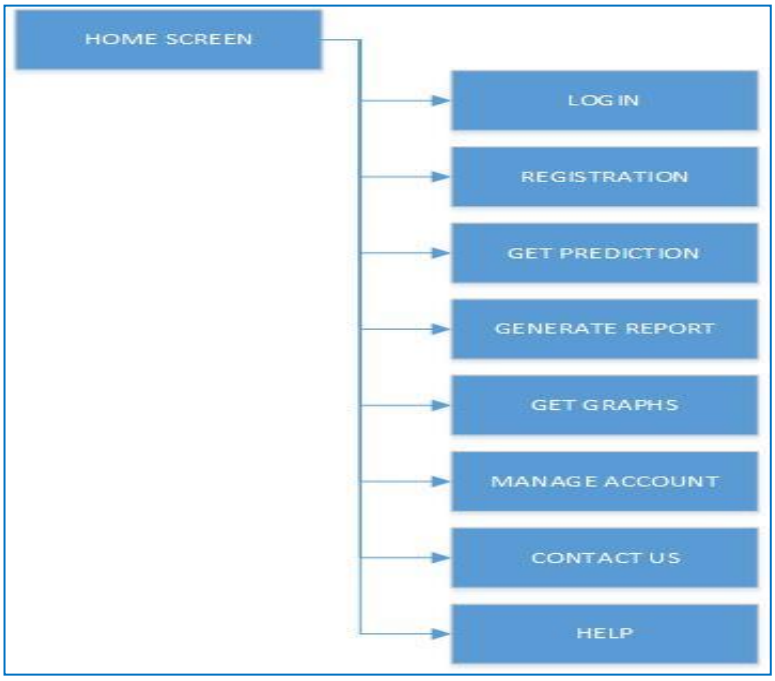

Figure 4-10 Web Pages

4.5.3 Client Side Flow Chart:

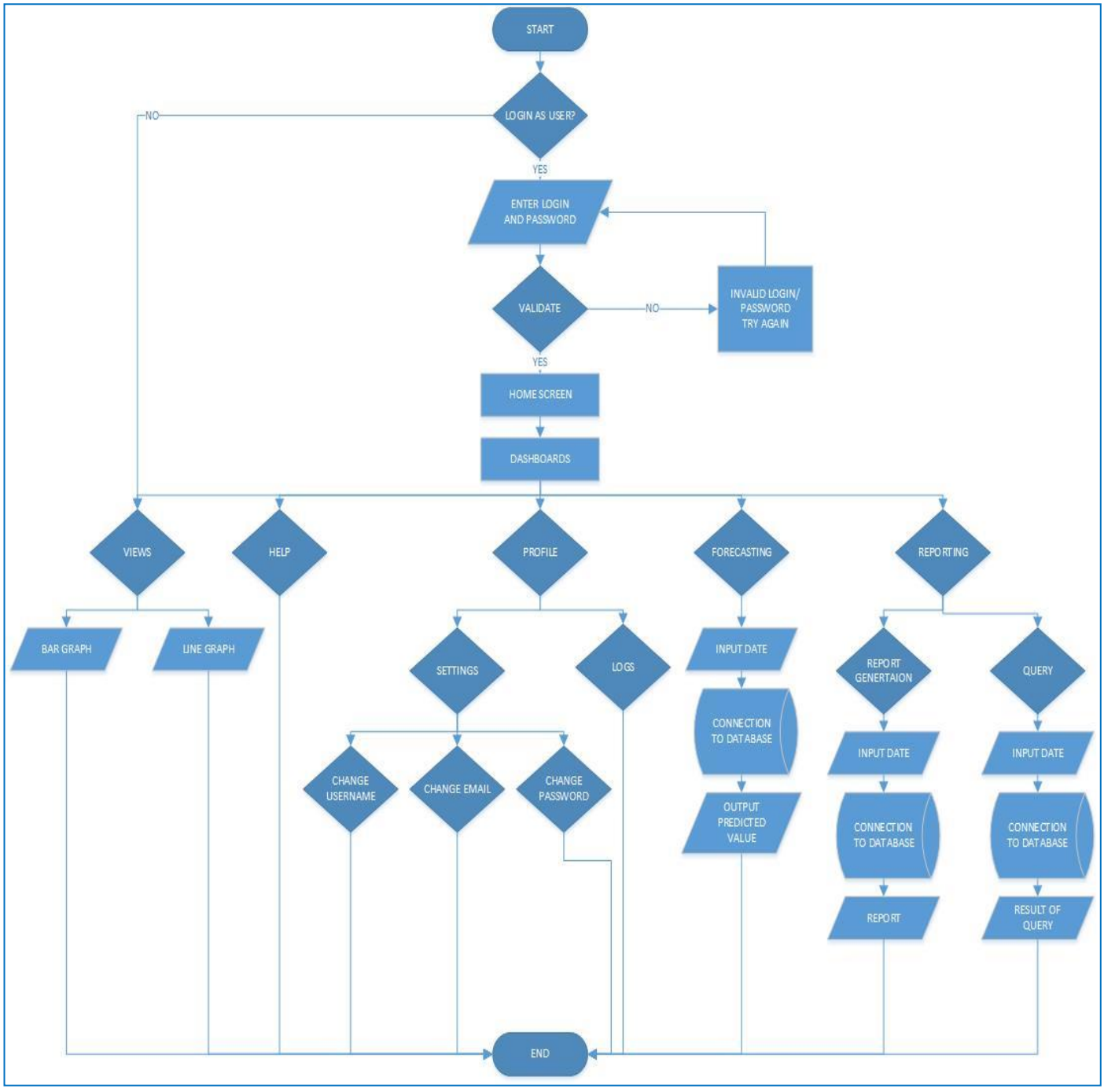

Figure 4-11 Client Side Flow Chart

4.5.4 Client Side Use Case Diagram:

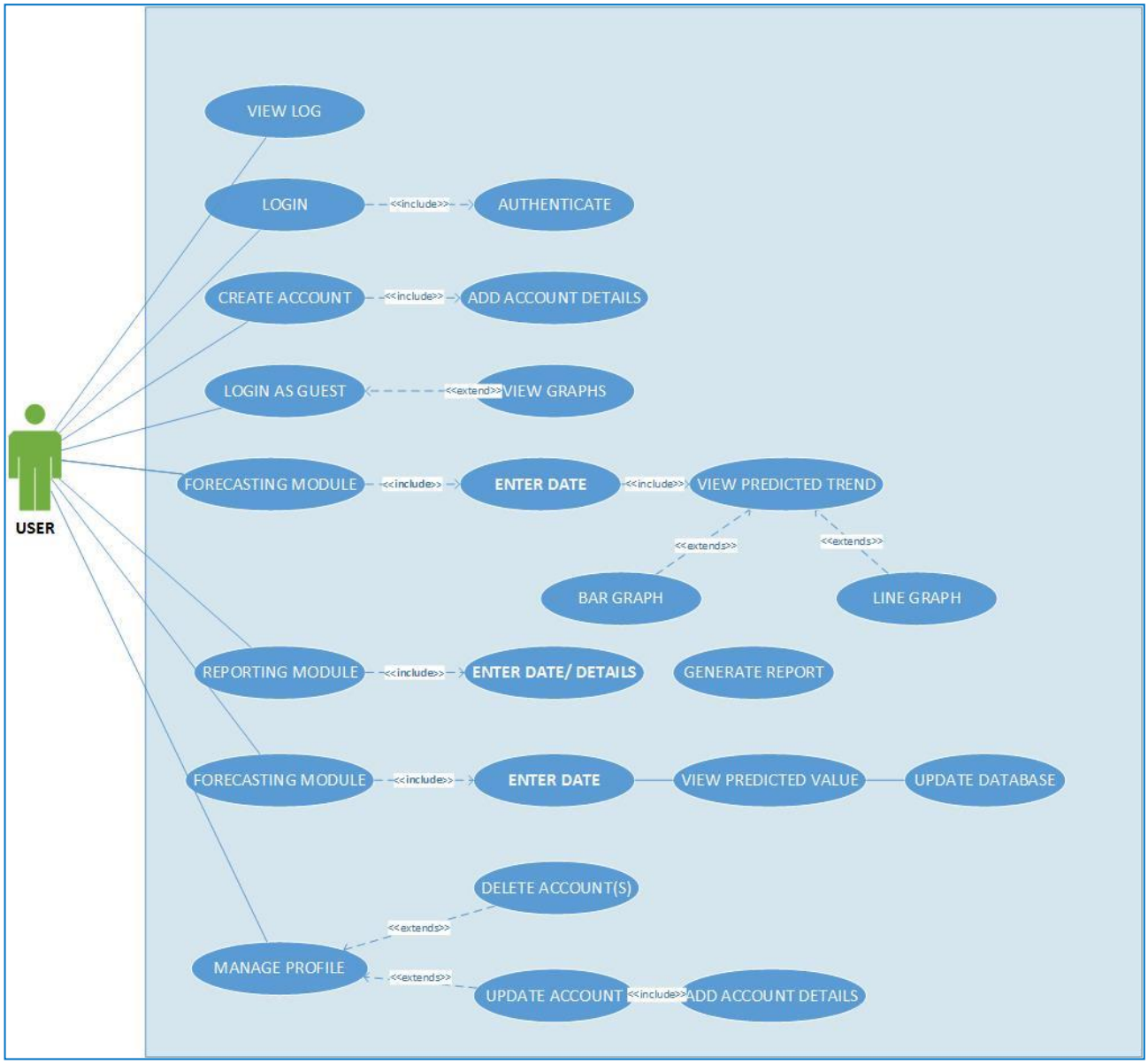

Figure 4-12 Client Side Use Case Diagram

4.6 Data Design

4.6.1 Data Model:

Dimensional data modeling is used for structuring the data warehouse. Dimensional Modeling comprises of one or more dimension table and fact tables. In our data warehouse there is only one fact table and every variable factor has its own dimension table. The dimension table has the data about the variable, which affects the oil price. The fact table stores the foreign keys of the dimension tables.

Dimensional modeling is a technique for conceptualizing and visualizing data models as a set of measures that are described by common aspects of the business. Dimensional modeling has two basic concepts:

i. **Facts:**
   A fact is a collection of related data items, consisting of measures.
ii. **Dimension:**
   The parameter over which we want to perform analysis of facts.

4.6.2 Data Dictionary:

4.6.2.1 User Database: (Stores data about user accounts)

Table Name	Field	Type	Null	Default
UserProfile	Id	Primary Key, int	Not Null	-
	Name	varchar(50)	Null	-
	Username	varchar(50)	Null	-
	Password	varchar(50)	Null	-
	Email	varchar(25)	Null	-
	ProfilePhoto	varchar(50)	Null	-

Table 4- 1 User Profile User Database

4.6.2.2 Staging Database: (Volatile data storage)

Table Name	Field	Type	Null	Default
TableCPI	Date	Primary Key, date	Not Null	-
	cpiuIndex	nchar(10)	Null	-
	cpiuYearOverYear	nchar(10)	Null	-
	cpiuGrowthRate	nchar(10)	Null	-
	cpiwIndex	nchar(10)	Null	-
	cpiwYearOverYear	nchar(10)	Null	-

Table 4- 2 CPI Staging Database

Table Name	Field	Type	Null	Default
TableOilPrice	Date	Primary Key,datetime	Not Null	-
	priceWTI	nchar(10)	Null	-

Table 4- 3 Oil Price Staging Database

Table Name	Field	Type	Null	Default
TableProduction	Date	Primary Key, datetime	Not Null	-
	Production	nchar(10)	Null	-

Table 4- 4 Production Staging Database

Table Name	Field	Type	Null	Default
TableStock	Date	Primary Key, datetime	Not Null	-
	Stock	nchar(10)	Null	-

Table 4- 5 Stock Staging Database

Table Name	Field	Type	Null	Default
TableSupply	Date	Primary Key, datetime	Not Null	-
	Supply	nchar(10)	Null	-

Table 4- 6 Supply Staging Database

4.6.2.3 Data-vault Database: (Non-Volatile data storage for forecasting purpose)

Table Name	Field	Type	Null	Default	Description
VaultFactTable	date	Primary Key, date	Not Null	-	primary key
	idCpi	Int	Null	-	Foreign key
	idStock	Int	Null	-	Foreign key
	idSupply	Int	Null	-	Foreign key
	idProduction	Int	Null	-	Foreign key
	idOilPrice	Int	Null	-	Foreign key

Table 4- 7 Fact Table Data-vault Database

Table Name	Field	Type	Null	Default	Description
VaultCPI	idCpi	int	Not Null	-	primary key
	dataCpiu	float	Null	-	
	dataCpiw	float	Null	-	Value of CPI

Table 4- 8 CPI Data-vault Database

Table Name	Field	Type	Null	Default	Description
VaultStock	idStock	Int	Not Null	-	primary key
	dataStock	Float	Null	-	Value of Stock

Table 4- 9 Stock Data-vault Database

Table Name	Field	Type	Null	Default	Description
VaultSupply	idSupply	Int	Not Null	-	primary key
	dataSupply	Float	Null	-	Value of Supply

Table 4- 10 Supply Data-vault Database

Table Name	Field	Type	Null	Default	Description
VaultProduction	idProduction	int	Not Null	-	primary key
	dataProduction	float	Null	-	Value of Production

Table 4- 11 Production Data-vault Database

Table Name	Field	Type	Null	Default	Description
VaultOilPrice	idOilPrice	Int	Not Null	-	primary key
	dataOilPriceWTI	float	Null	-	Value of Oil Price

Table 4- 12 Oil Price Data-vault Database

4.7 Entity Relation Diagram:

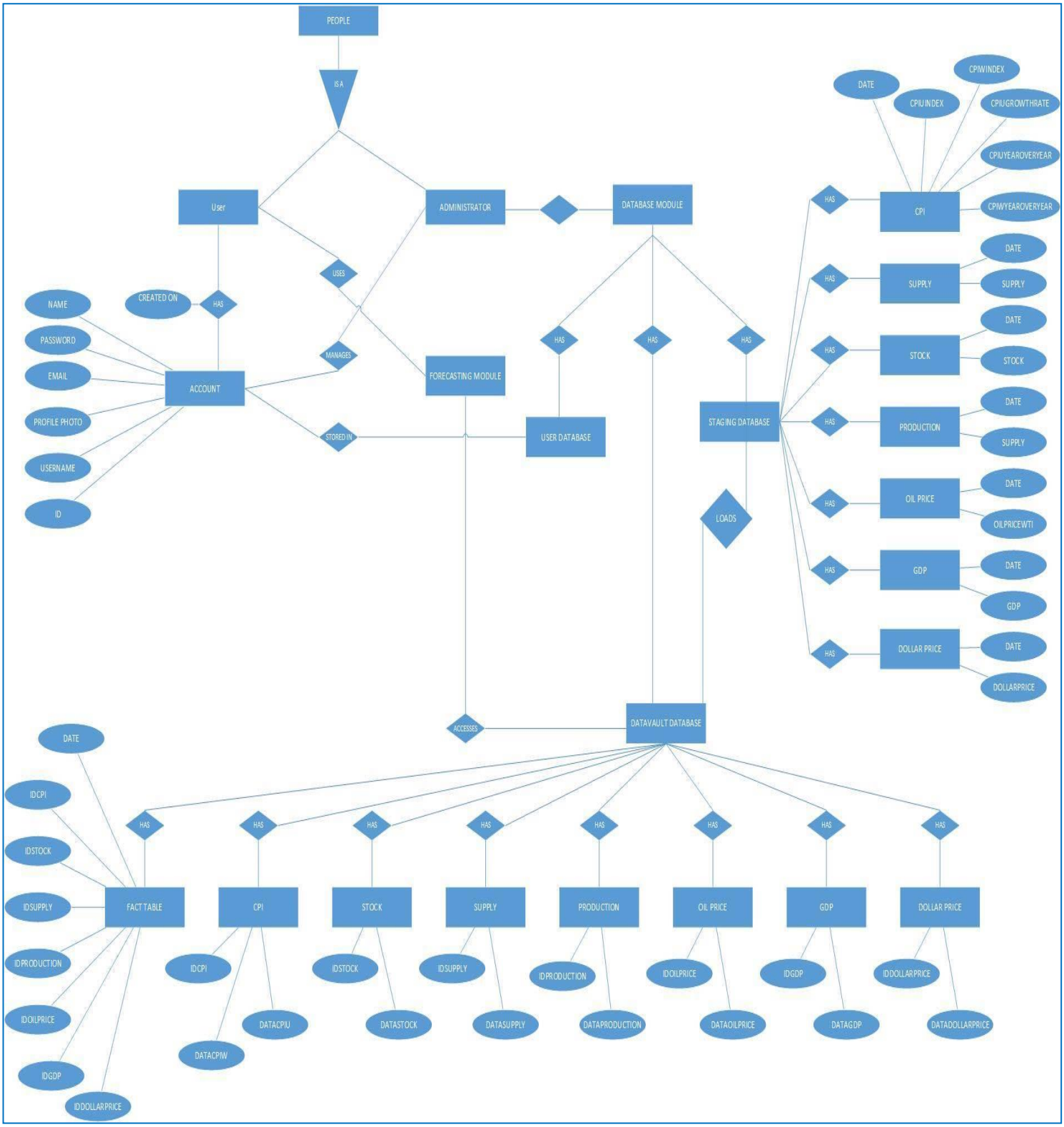

Figure 4-13 Entity Relation Diagram

4.8 UML Diagrams

4.9 Use Case Diagrams:

4.9.1 Forecast Oil Price Value:

4.9.1.1 Use Case Description:

Use Case Name	**Forecast Oil Price Value**
Actor(s)	User
Description	The user requests the System to generate/predict Oil price(s) of a particular day, month, or year.
Normal Course	i. This use case is initiated when the user wants to predict Oil price(s). ii. The System queries the user to enter the date. iii. The user enters the day, date and year and presses the generate button. iv. The System shows the desired result to the user.
Alternate Course	The user leaves the date fields empty and presses the generate button, the System prompts an error message.
Pre-Condition	The user credentials have been verified.
Post-Condition	The new predicted value is stored in the database for future predictions.
Assumptions	i. Online Server is running. ii. Database is connected to the Server. iii. Browser is compatible to render ASP.Net page(s)

Table 4- 13 Forecast Oil Price Value Description

4.9.1.2 Use Case Diagram:

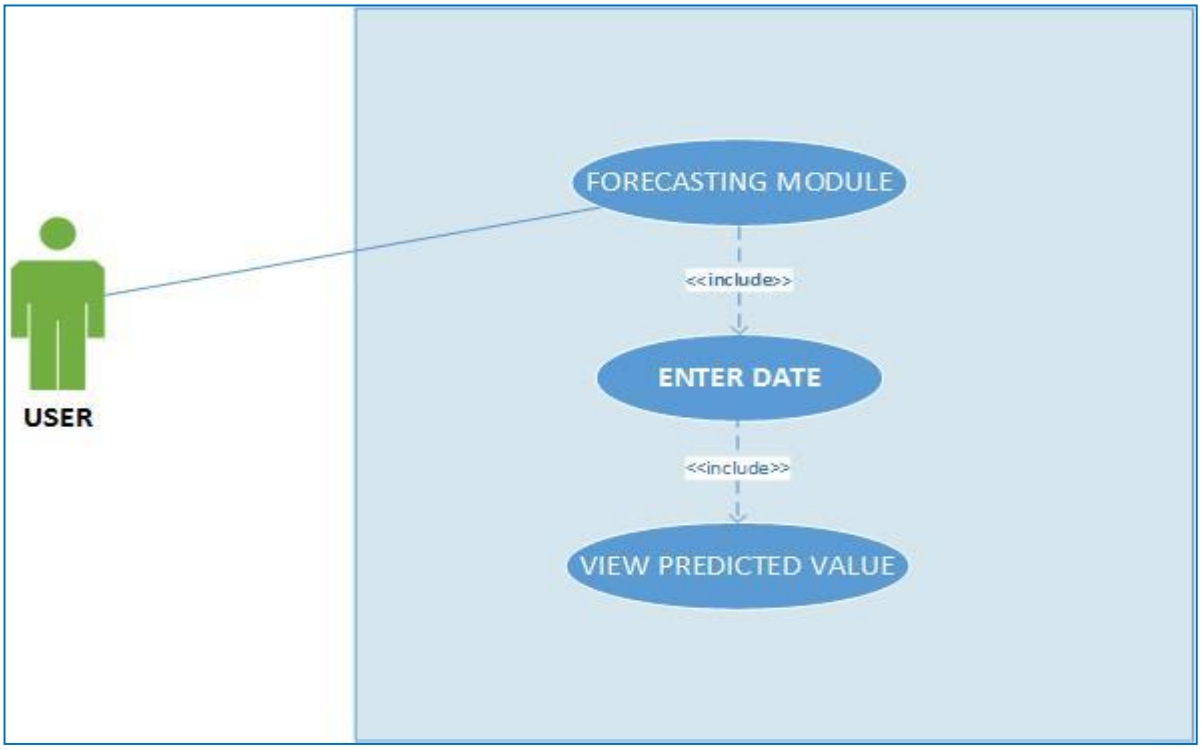

Figure 4-14 Forecast Oil Price Value Use Case Diagram

4.9.2 Maintain Database with newly Predicted Value(s):

4.9.2.1 Use Case Description:

Use Case Name	**Maintain Database with newly Predicted Value(s)**
Actor(s)	User, Secondary Actor: Database
Description	When the user requested the System to predict the price and the System has predicted an outcome, that outcome will also be stored in the database for future use.
Normal Course	i. This use case is initiated when a user has requested to predict Oil price on some future date. ii. The System has displayed the Predicted the value. iii. The Systems now stores this predicted value in the database for future use.
Alternate Course	-
Pre-Condition	The user credentials have been verified.
Post-Condition	The Database has been updated.
Assumptions	i. Online Server is running. ii. Database is connected to the Server. iii. Browser is compatible to render ASP.Net page(s).

Table 4- 14 Maintain Database with newly Predicted Value(s) Description

4.9.2.2 Use Case Diagram:

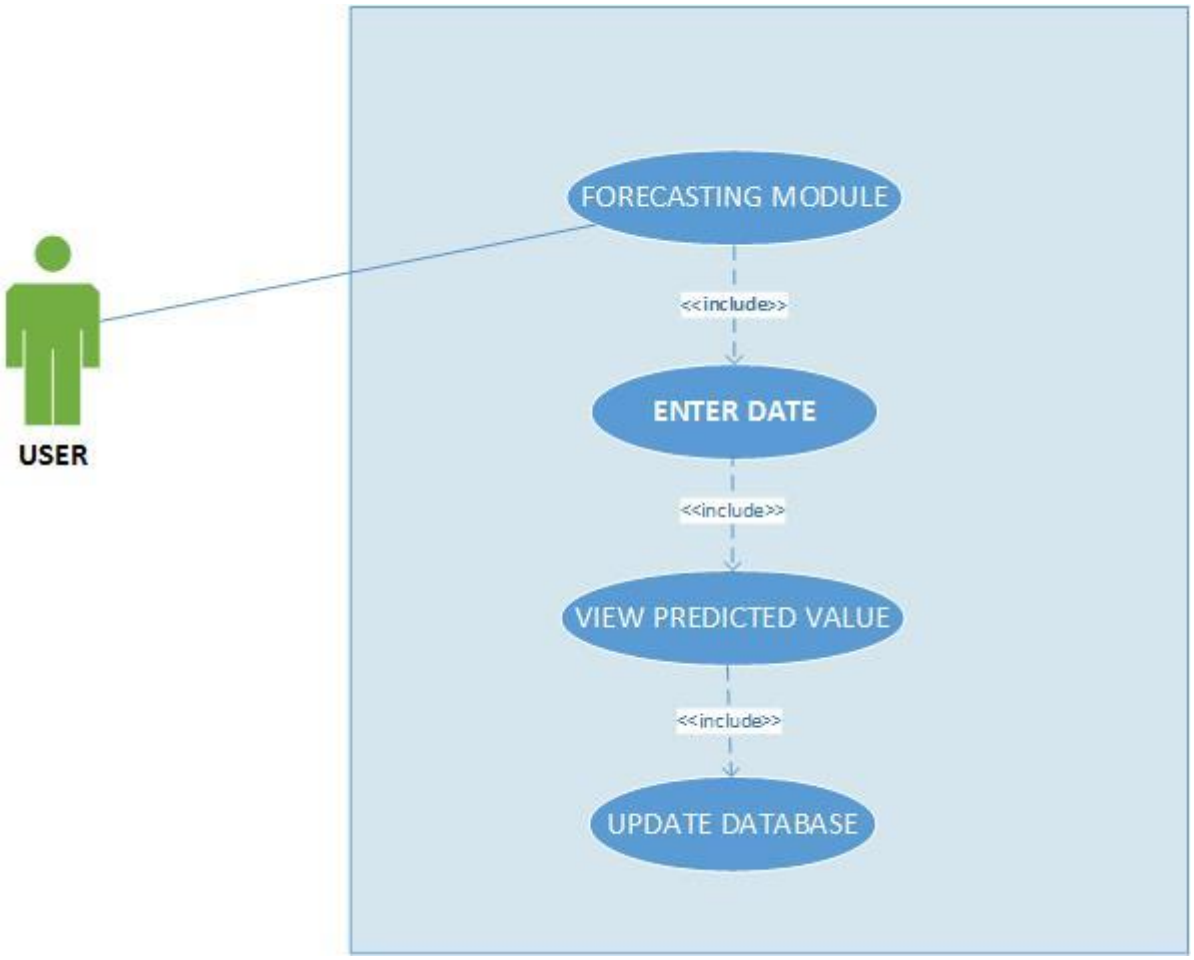

Figure 4-15 Maintain Database with newly Predicted Value(s) Use Case Diagram

4.9.3 User Login:

4.9.3.1 Use Case Description:

Use Case Name	**User Login**
Actor(s)	User, Secondary Actor: Database
Description	User is required to login in order to use the features of the System.
Normal Course	i. This use case is initiated when the user wants to access the Web Application. ii. The user accesses the login page. iii. The System asks the user to enter his username and password. iv. The user enters his username and password and presses submit. v. The System verifies the credentials from the server and allows the user access. vi. The user accesses the login page and logs in as guest. vii. The System allows the user to view general graphs.
Alternate Course	i. The System verifies the credentials, the login details are wrong, the System shall prompt an error message. ii. If the user does not have an account, the System allows the user to create an account and asks him for following information: i. Name of the User ii. Email of the User iii. Username iv. Password.
Pre-Condition	-
Post-Condition	The Database has been updated.
Assumptions	i. Online Server is running. ii. Database is connected to the Server. iii. Browser is compatible to render ASP.Net page(s).

Table 4- 15 User Login Description

4.9.3.2 Use Case Diagram:

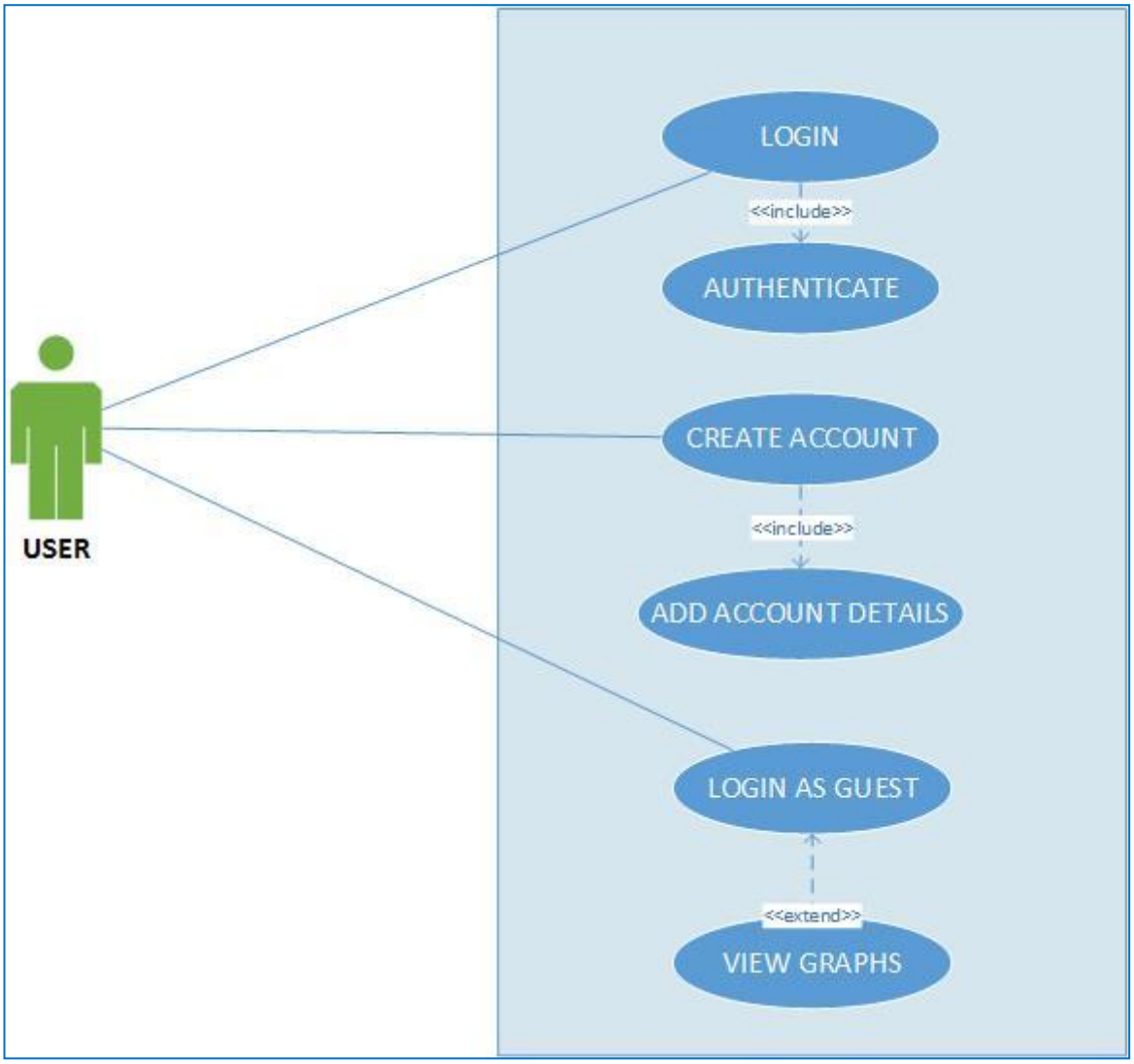

Figure 4-16 User Login Use Case Diagram

4.9.4 View Graphs of trends in Oil Price:

4.9.4.1 Use Case Description:

Use Case Name	**View graphs of trends in Oil Price**
Actor(s)	User
Description	The user requests the system to generate graph(s) for viewing Oil price trends of a particular year or range.
Normal Course	i. This use case is initiated when the user wants to predict Oil price(s). ii. The system queries the user to enter the year or range. iii. The user enters the year/ range. iv. The system generates graphs of the whole year with relevant axes mentioned
Alternate Course	The user leaves the date fields empty and presses the generate button, the system prompts an error message.
Pre-Condition	The user credentials have been verified.
Post-Condition	i. The new predicted value(s) is stored in the database for future predictions. ii. The required graph is displayed.
Assumptions	i. Online Server is running. ii. Database is connected to the Server. iii. Browser is compatible to render ASP.Net page(s).

Table 4- 16 View graphs of trends in Oil Price Description

4.9.4.2 Use Case Diagram:

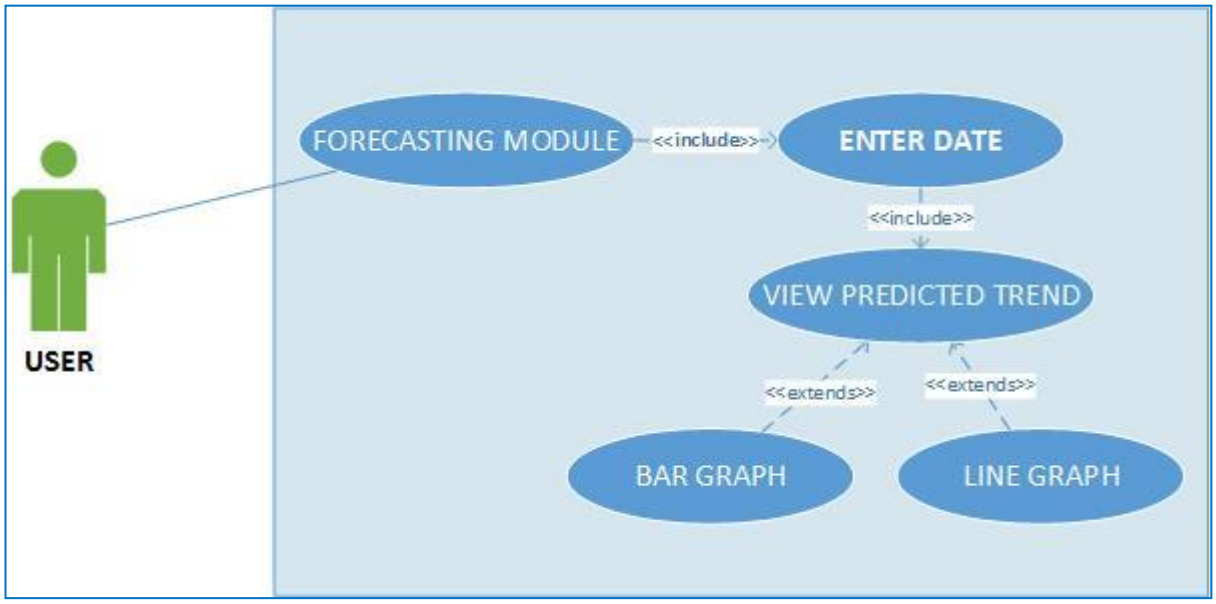

Figure 4-17 View graphs of trends in Oil Price Use Case Diagram

4.9.5 Administrator Role on Server Side:

4.9.5.1 Use Case Description:

Use Case Name	Administrator Role on Server Side
Actor(s)	Admin, Secondary Actor: Database
Description	Administrator has the right to generate log reports of any particular time period /person and to create accounts of users.
Normal Course	i. This use case is initiated when the admin accesses the login page on the Server side. ii. The System asks the admin to enter Login details. iii. The admin enters his username and password and presses submit. iv. The System verifies the credentials from the Server and allows the admin access. v. The System allows admin to: i. Manage Account. ii. Manage Database vi. If the admin selects Account Creation, the System asks him for Account information: vii. If the admin selects Update Database, the System allows admin to update the following databases: i. Staging Database. ii. DataVault Database. iii. User Database. viii. Database, the System allows admin to delete following databases: i. Staging Database. ii. DataVault Database. iii. User Database.
Alternate Course	The System verifies the credentials, the login details are wrong, the System shall prompt an error message.
Pre-Condition	i. Online Server is running. ii. Database is connected to the Server.
Post-Condition	i. Successful Completion: The system will confirm that the particular task has been performed and results if required are saved in the database. ii. Failure Condition: The changes will not be stored and the system will display a message accordingly.

Table 4- 17 Administrator Role on Server Side Description

4.9.5.2 Use Case Diagram:

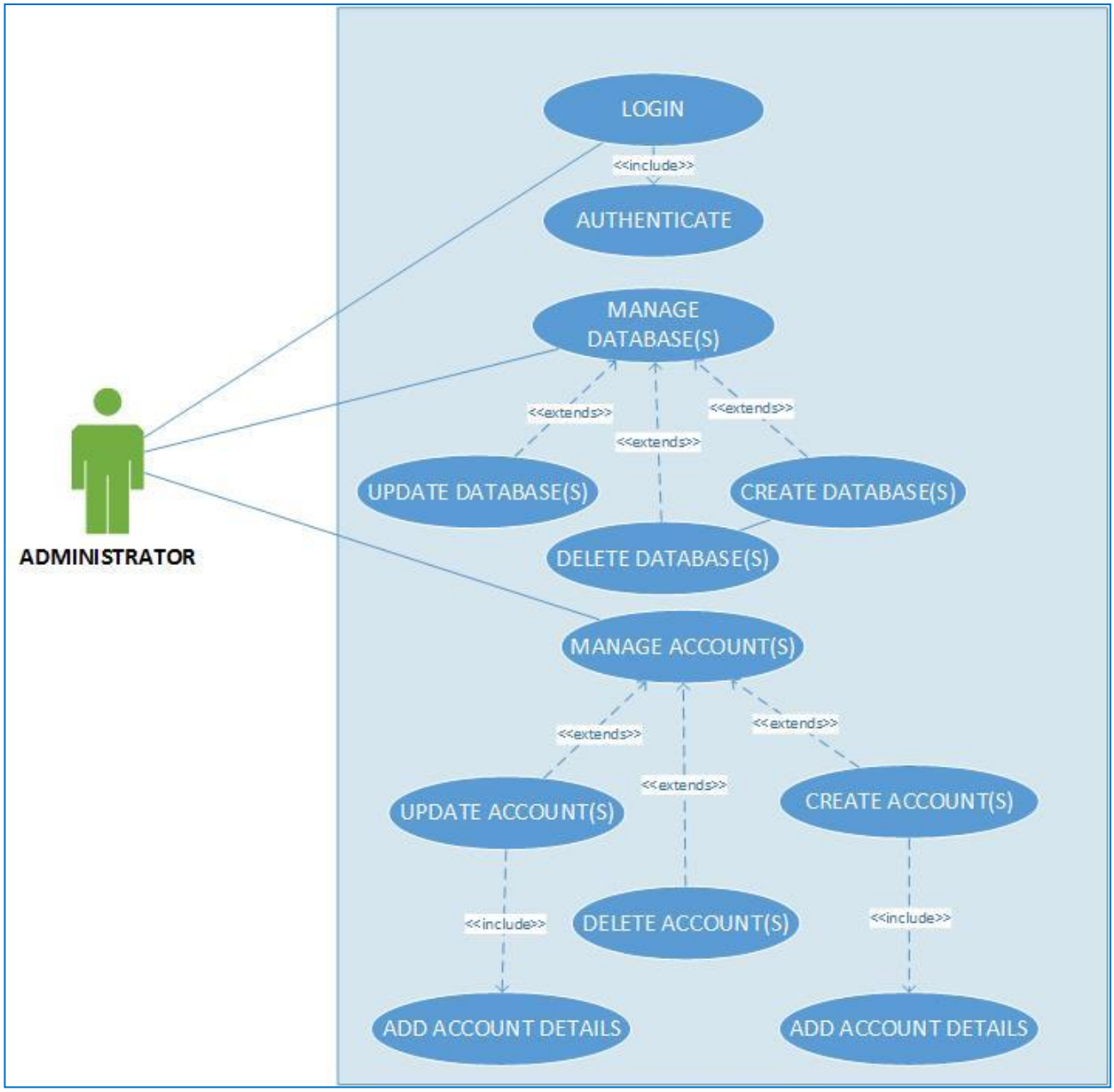

Figure 4-18 Administrator Role on Server Side Use Case Diagram

4.9.6 Report Generation:

4.9.6.1 Use Case Description:

Use Case Name	Report Generation
Actor(s)	User
Description	The user requests the System to generate report.
Normal Course	i. This use case is initiated when the user wants to generate report. ii. The System queries the user to enter the details. iii. The user enters the details and presses the generate button. iv. The System shows the desired result to the user.
Alternate Course	The user does not enter the details and presses the generate button, the System prompts an error message.
Pre-Condition	The user credentials have been verified.
Post-Condition	The System shows the desired result to the user.
Assumptions	i. Online Server is running. ii. Database is connected to the Server. iii. Browser is compatible to render ASP.Net page(s)

Table 4- 18 Report Generation Description

4.9.6.2 Use Case Diagram:

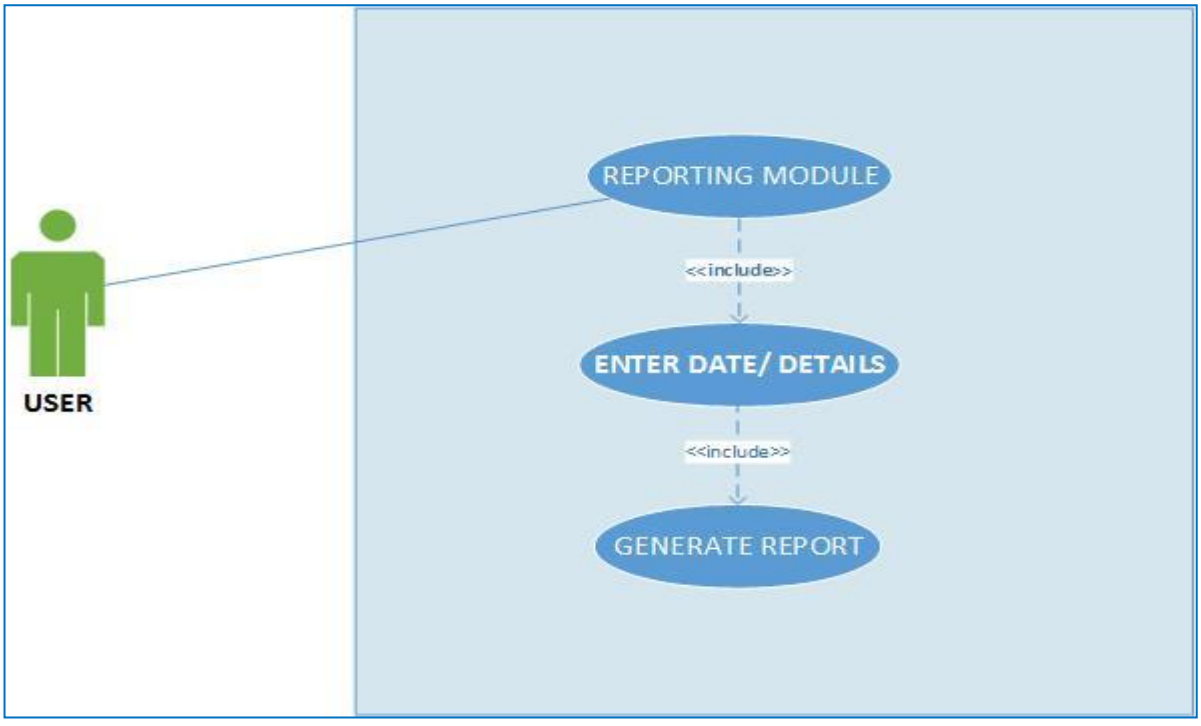

Figure 4-19 Report Generation Use Case Diagram

4.9.7 Manage Profile

4.9.7.1 Use Case Description

Use Case Name	**Manage Profile**
Actor(s)	User
Description	The user can edit the account details and view log.
Normal Course	i. The user is logged in successfully and wants to edit the account details. ii. The System allows user to: Manage Account. i. Update Account. ii. Delete Account. iii. The user selects Update Account. iv. The System allows the user to edit the account information: i. Username ii. Email iii. Password. v. The user selects Delete Account. vi. The System deletes the Account. vii. The user selects to view log. viii. The System displays log.
Alternate Course	-
Pre-Condition	The user credentials have been verified.
Post-Condition	Database is updated accordingly.
Assumptions	i. Online Server is running. ii. Database is connected to the Server. iii. Browser is compatible to render ASP.Net page(s)

Table 4- 19 Manage Profile Description

4.9.7.2 Use Case Diagram

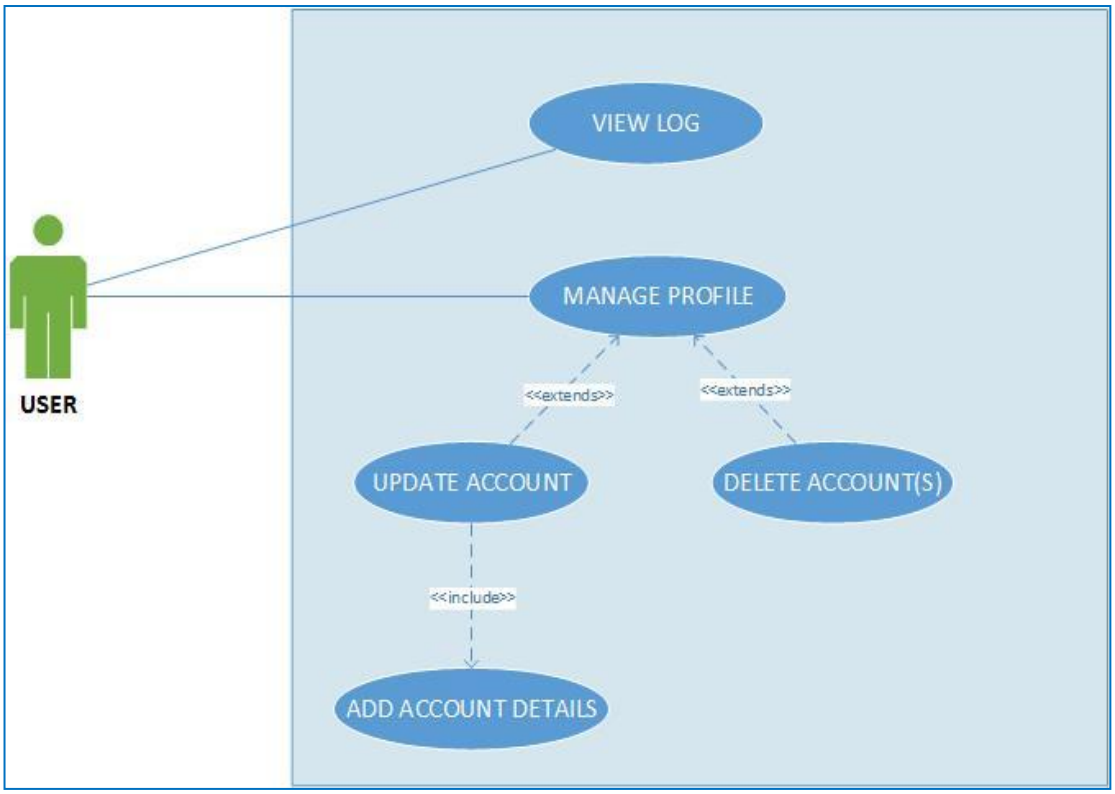

Figure 4-20 Manage Profile Use Case Diagram

4.10 Sequence Diagrams

4.10.1 Forecast Oil Price Value

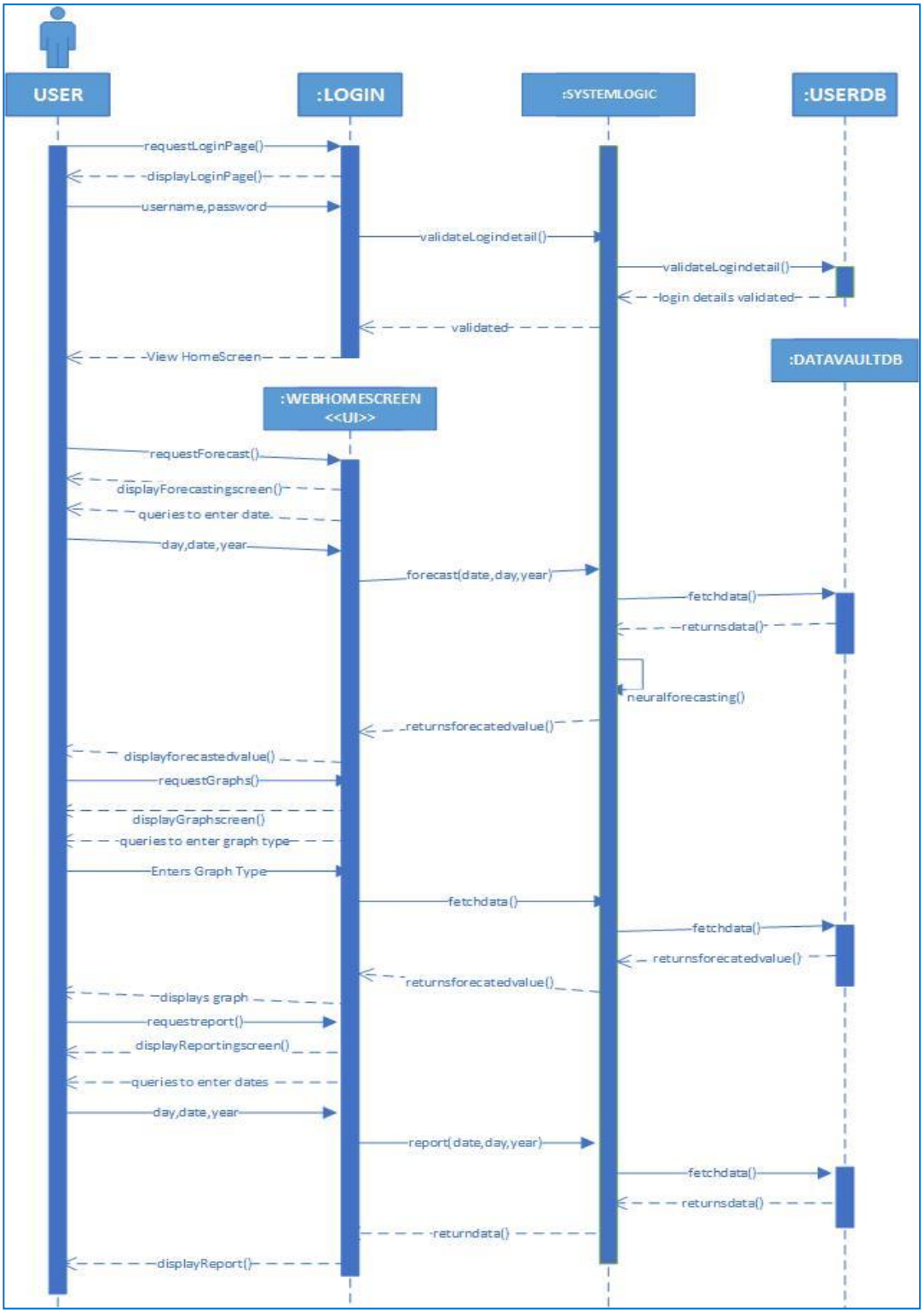

Figure 4-21 Forecast Oil Price Value Sequence Diagram

4.10.2 Manage Profile

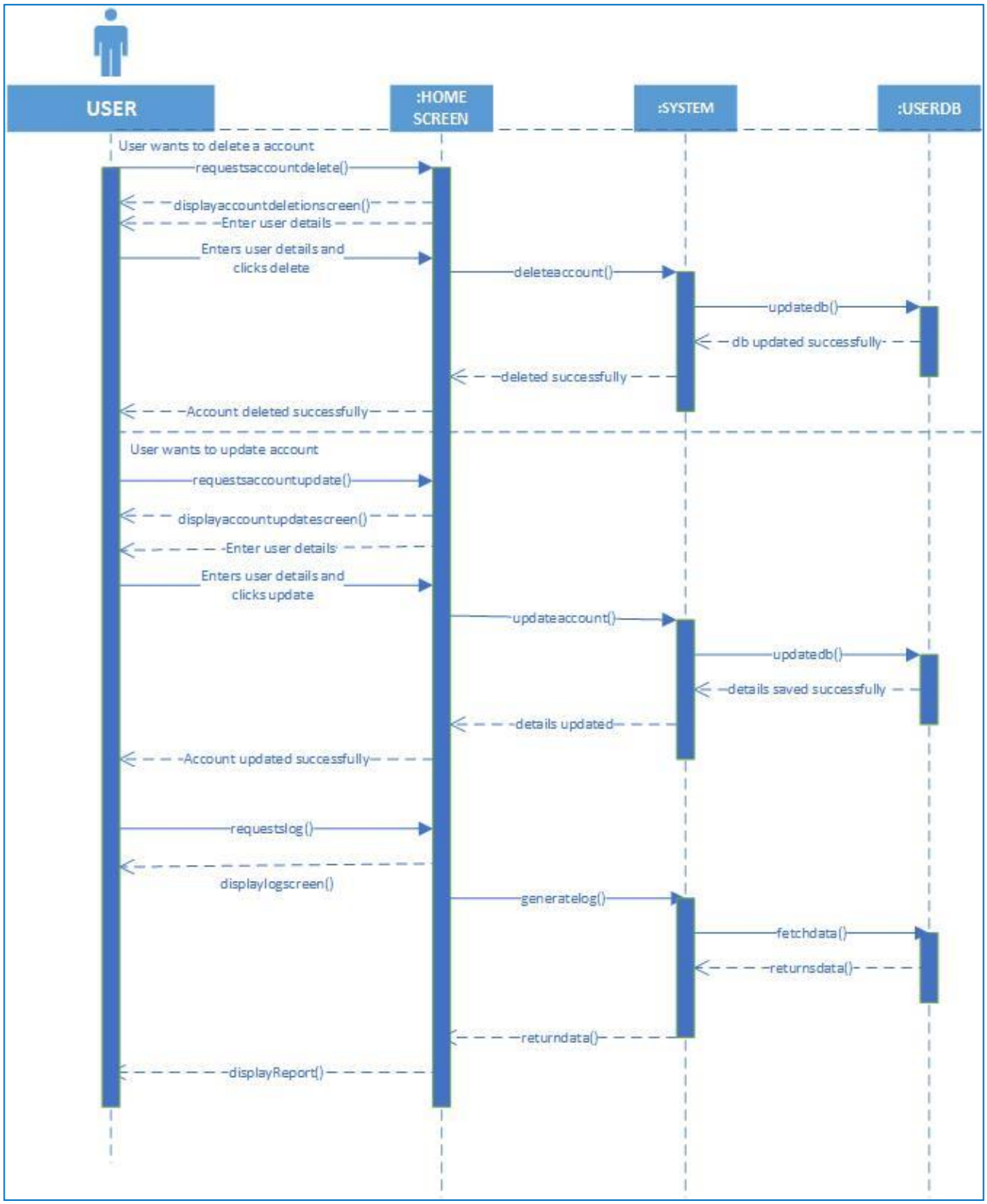

Figure 4-22 Manage Profile Sequence Diagram

4.10.3 Administrator Role on Server Side

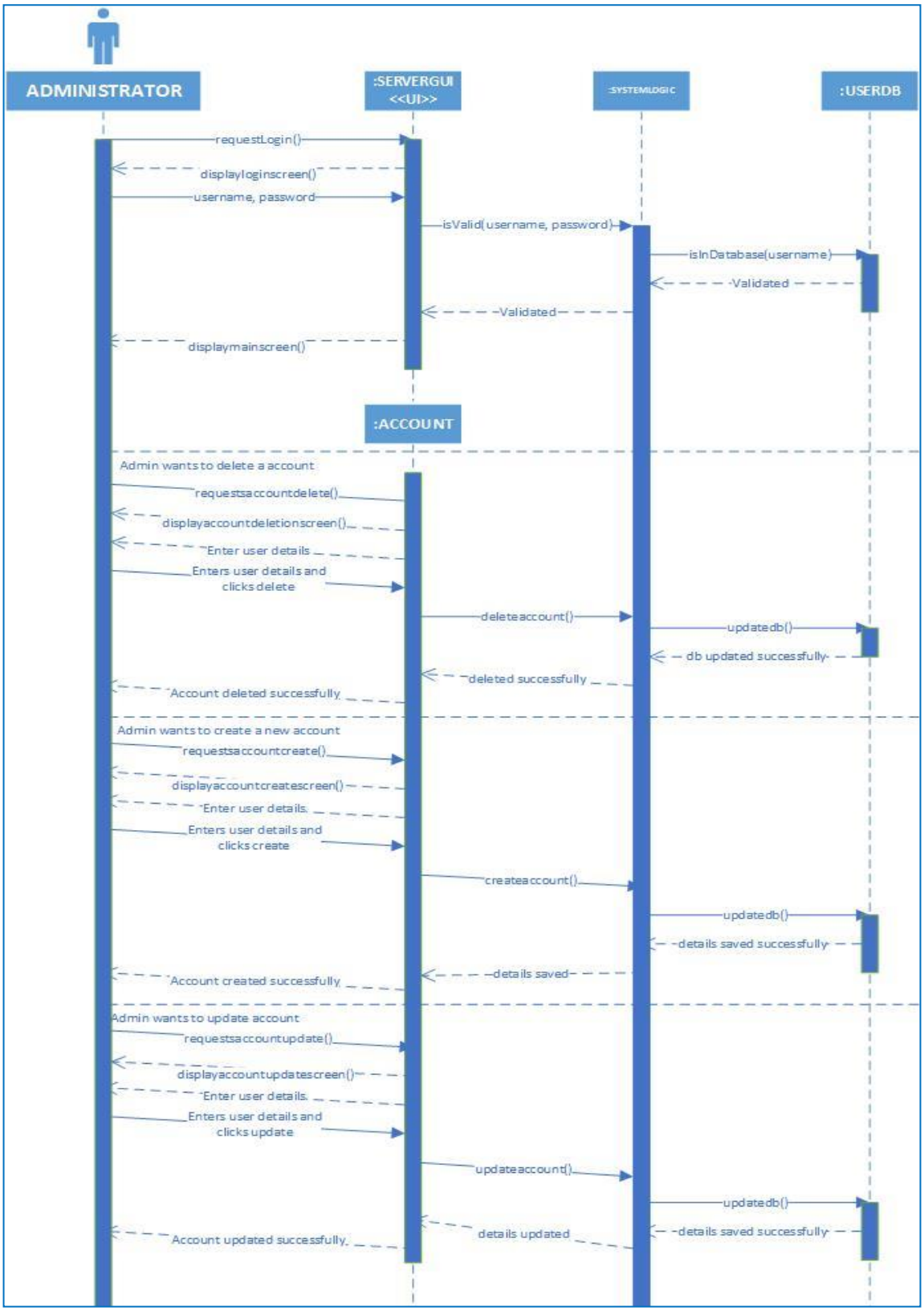

Figure 4-23 Administrator Role on Server Side Sequence Diagram

4.11 Activity Diagrams

4.11.1 Forecast Oil Price Value/Generate Graph and Report

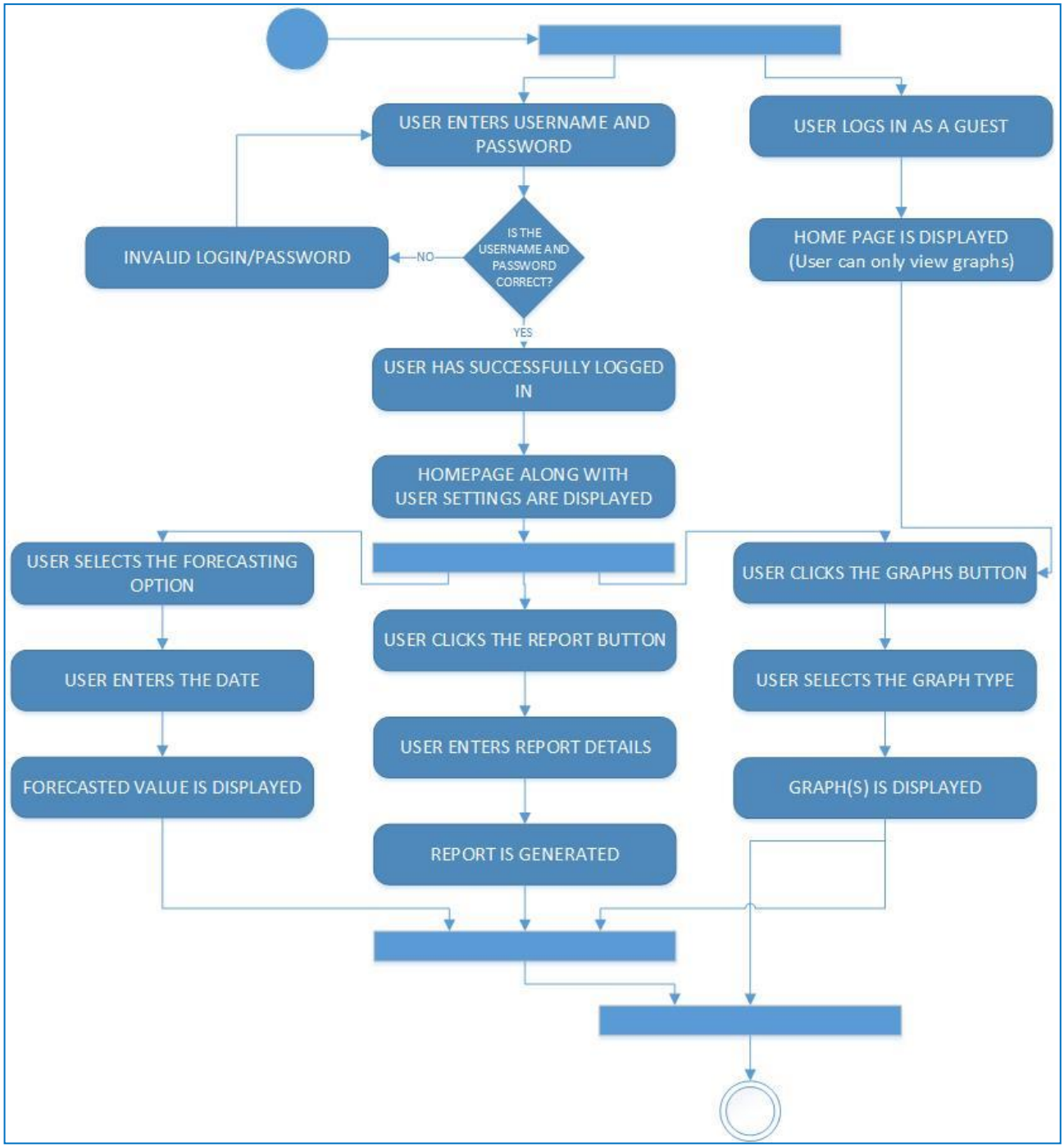

Figure 4-24 Forecast Oil Price Value/Generate Graph and Report Activity Diagram

4.11.2 Administrator Role on Server Side

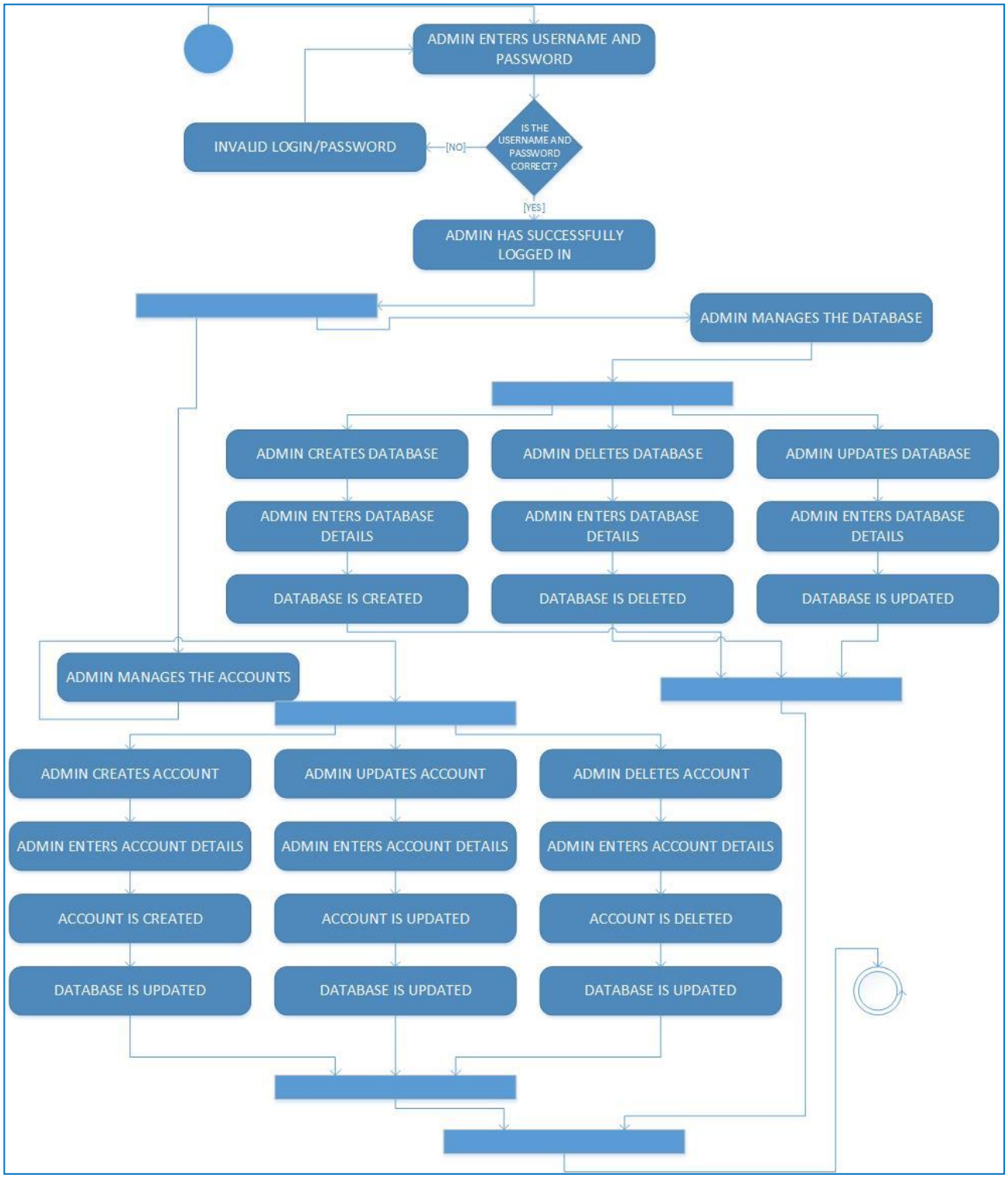

Figure 4-25 Administrator Role on Server Side Activity Diagram

4.12 State Transition Diagrams

4.12.1 Forecast Oil Price Value and Generate Report

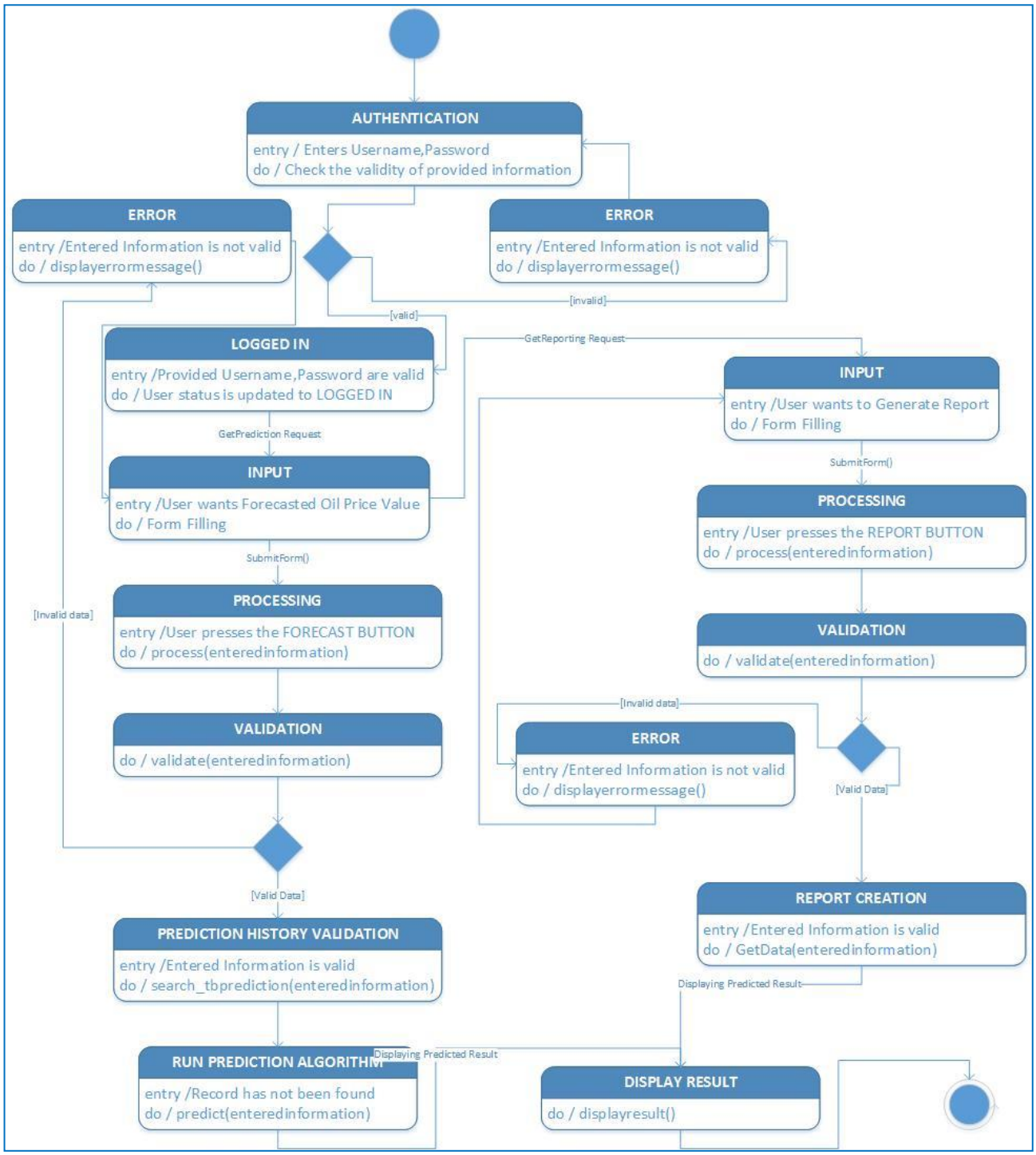

Figure 4-26 Forecast Oil Price Value and Generate Report State Transition Diagram

4.12.2 Generate Graphs

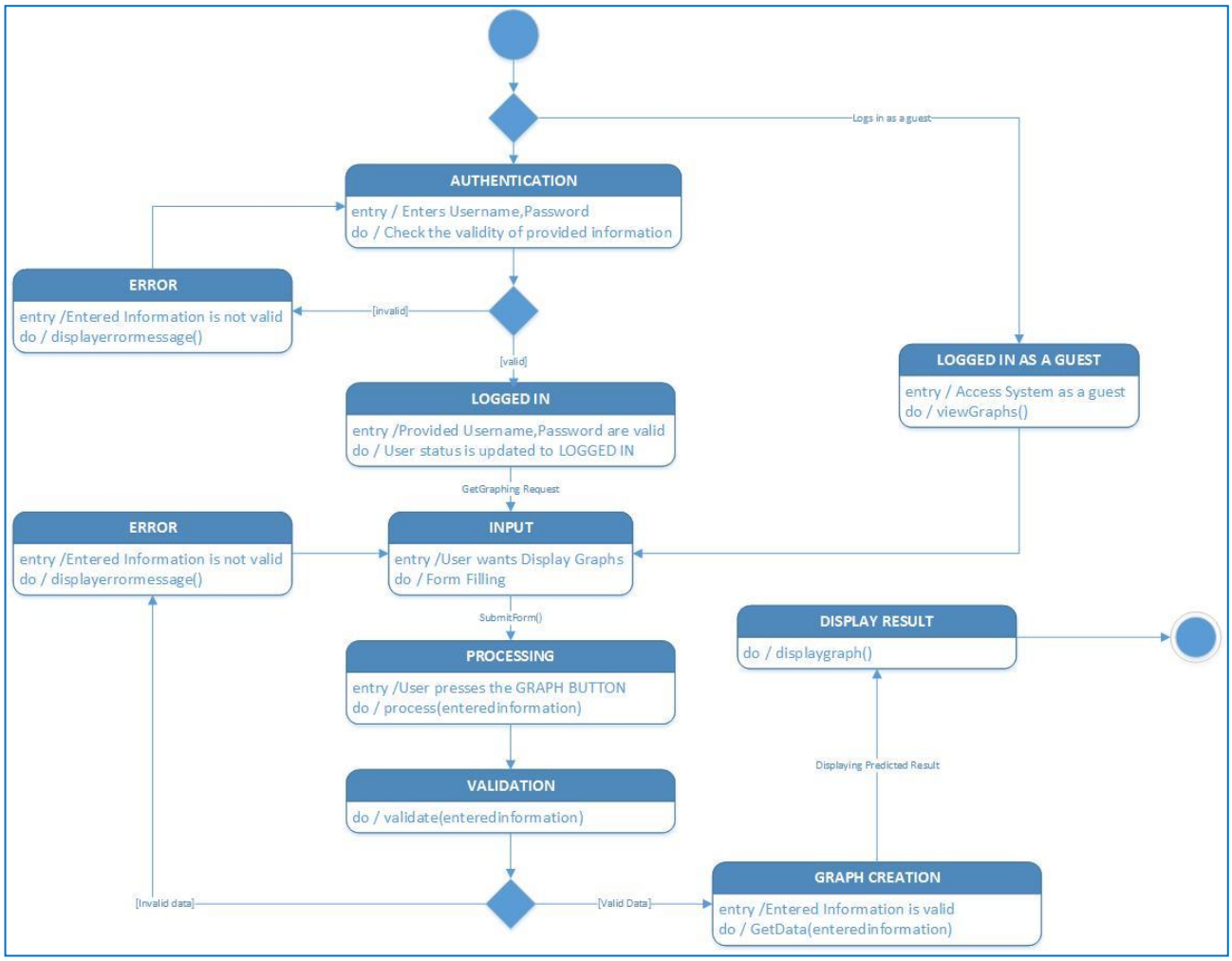

Figure 4-27 Generate Graph State Transition Diagram

4.13 Data Flow Diagram

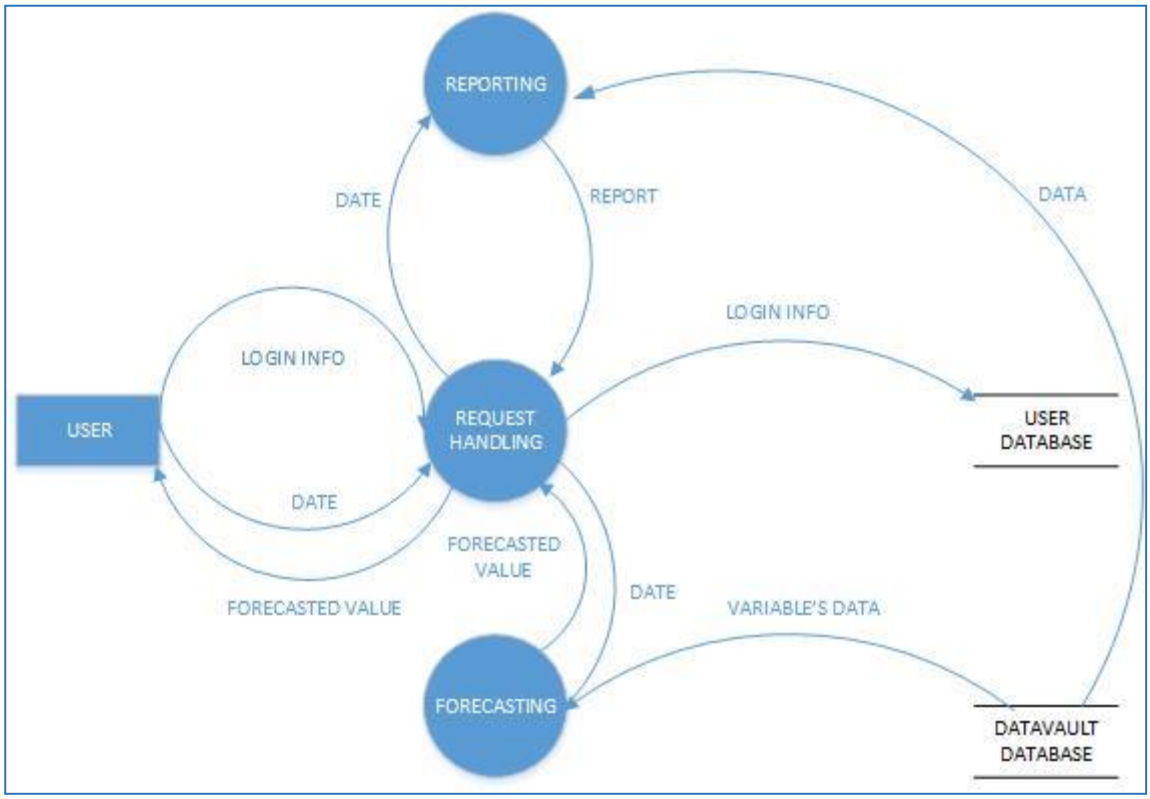

Figure 4-28 Data Flow Diagram

CHAPTER 5
IMPLEMENTATION

5. Implementation

5.1 Technology Used

1.1.1. Programming language

i. ASP.Net Framework 4.5 is used as programming language to develop Web pages.
ii. Java Script is used as client side scripting language.
iii. AJAX Control Toolkit is used to enhance client side features of website.
iv. LINQ to SQL provides a runtime infrastructure for managing relational data as objects without losing the ability to query.
v. AmCharts jquery library is used to create interactive graphs on client side.

1.1.2. Software Resources

i. Microsoft Visual Studio 2012
ii. Mozilla Firefox, Google Chrome etc.
iii. Microsoft SQL Server 2012

1.1.3. Database

Databases were developed and managed in Microsoft SQL.

1.1.4. Operating Environment

The main component of the Oil Price Forecasting Tool project is the Web application through which the user will give input to the system. The application is not resource or graphics-intensive, so there are no practical hardware constraints. The application will rely on several functionalities built into ASP. Net's Web Application Programming

Interface (Web API), so ensuring appropriate usage of the API will be a major concern. Beyond that, the application is a self-contained unit.

The application will however, frequently interact with the Server, where the database will be stored using SQL and will be interfaced with a wrapper written in ASP. Net.

5.2 Implementation Work

The main three modules that come under project scope are:

iv. **Historical Data (Data Collection):** It includes extracting the required information and analyze data from bulk of raw data and creating Data warehouse using this data.

Once the data has been collected, Data warehouse is created which stores all the historical data.

v. **Prediction Algorithm (Analysis):** It includes predicting Oil Price trends using Artificial Neural Networks and Time Series Data Mining Algorithm.

vi. **Results of Prediction (Presentation):** It includes generating Graphs and Report of analysis of prediction results.

5.2.1 Historical Data (Data Collection)

The web application maintains a data warehouse for storing all the data required by the forecasting tool for the prediction of oil prices. This data warehouse needs to be loaded regularly so that it can serve the purpose of facilitating data analysis. For this purpose, data from various sources is extracted and copied in to the Staging Database of the data warehouse. The main challenge is to integrate, rearrange and consolidate large amount of data, thereby providing a unified information base for business intelligence. The process of extracting data from sources and bringing it into data warehouse is commonly known as ETL.

i. Extraction of Data:

During extraction, the desired data is identified and extracted from many different online sources, including applications. The data is downloaded in the form of Excel files and is loaded into the Staging Database.

Data is collected from following online sources, in .xsls format, at the server side.

i. **Oil Prices:**

http://www.eia.gov/dnav/pet/hist_xls/RWTCd.xls

ii. **Production:**

http://tonto.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=WCRFPUS2&f=W

iii. **Stock:**

http://tonto.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=MTTSTUS1&f=M

iv. **Supply:**

http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=MTTUPUS1&f=M

v. **CPI:**

http://research.stlouisfed.org/fred2/series/CPIAUCSL/downloaddata

http://www.seattle.gov/financedepartment/cpi/historical.htm

ii. Transformation of Data:

Once the data is loaded into the Staging Database, cleansing is performed on the data i.e. data is normalized and rearranged. Staging database is a volatile database, once the data is loaded onto the data-vault, the data of staging database is deleted.

iii. Loading of Data:

After the cleansing of data, the data is loaded into the Data-vault Database. This is a non-volatile (permanent) data storage. From here the data is fetched for Neural Network training, Forecasting, Reporting and generating Graphs.

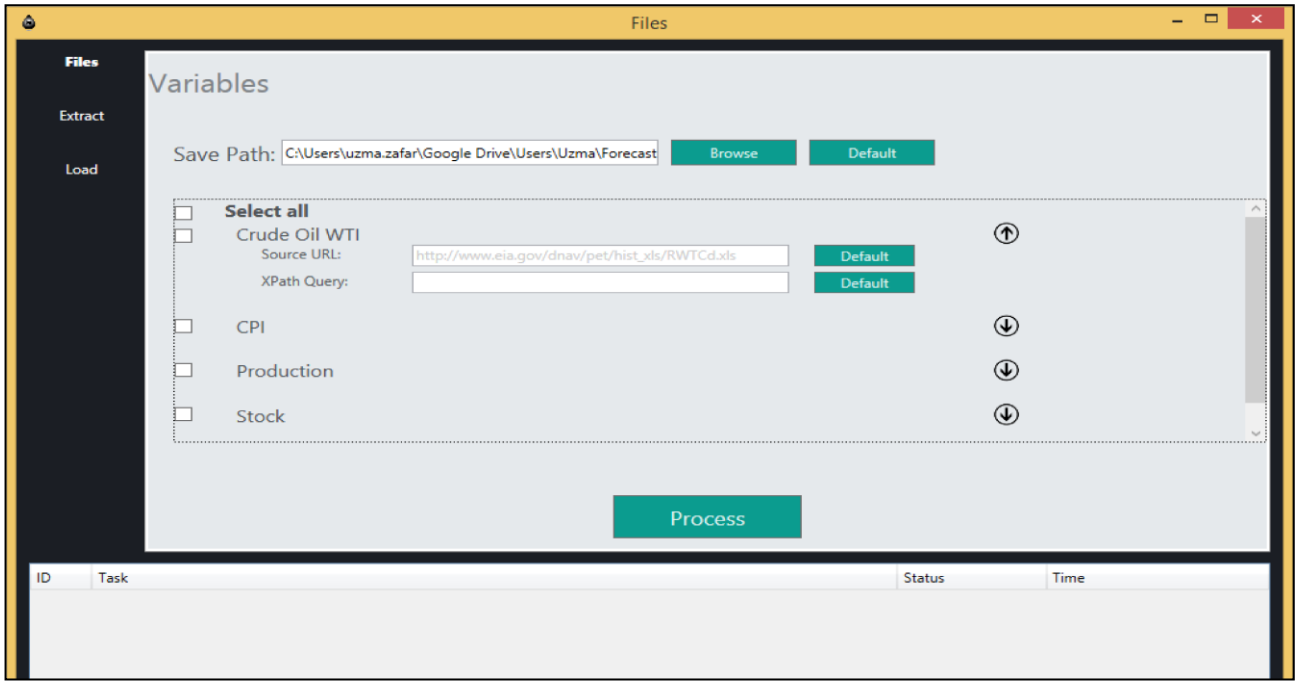

Figure 5-1 ETL Extract Screen

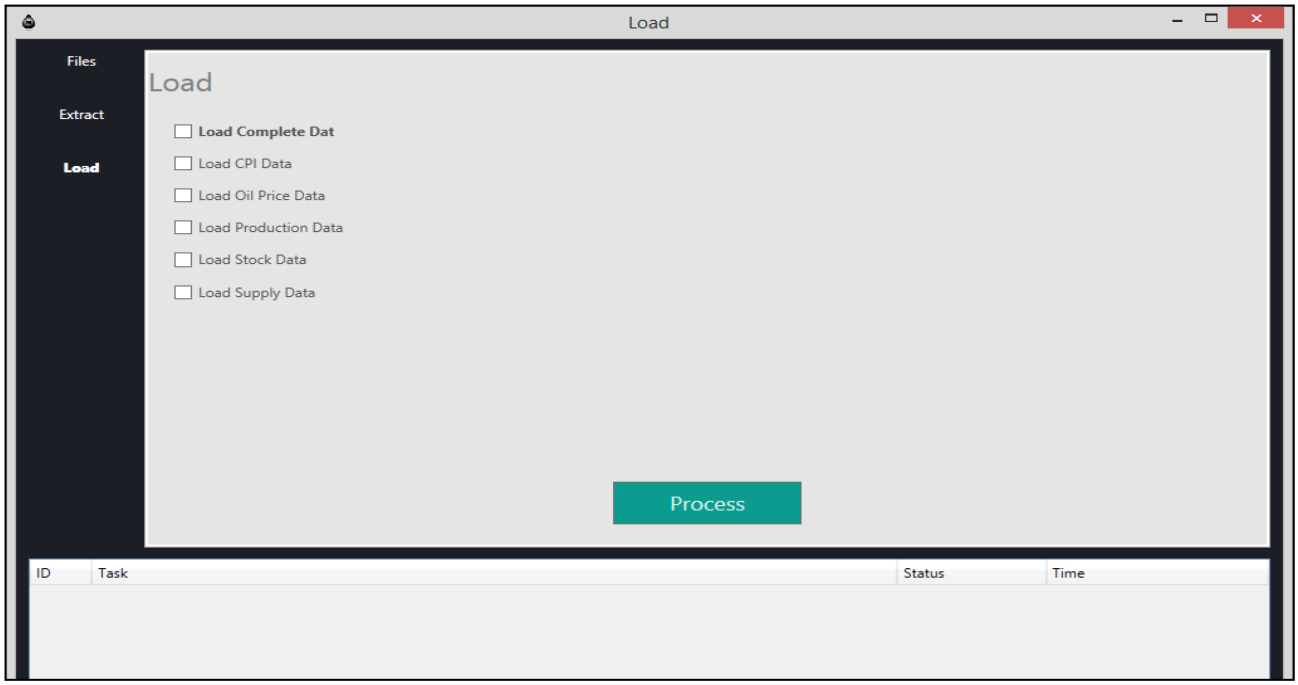

Figure 5-2 ETL Load Screen

5.2.2 Prediction Algorithm (Analysis)

The Forecasting Tool is implemented using the concepts of Artificial Neural Networks. It uses 80% of the data of data warehouse for the training of the network, which uses Back Propagation Algorithm. 20% of the data is used for forecasting the oil prices.

The basic steps in a Forecasting Process are:

i. Identify the Goal of the Forecast
ii. Establish a Time Horizon
iii. Collect Data
iv. Select a Forecasting Technique
v. Conduct the Forecast
vi. Monitor Accuracy

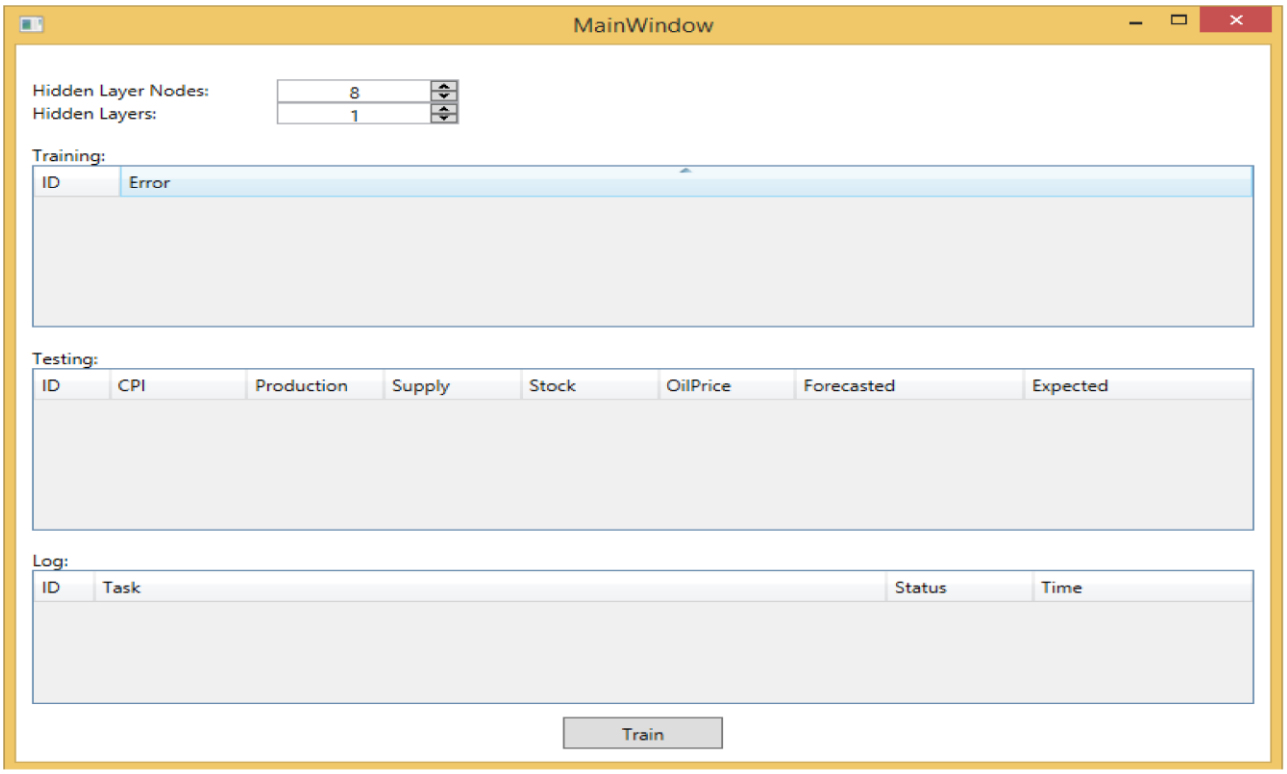

Figure 5-3 Forecasting Module Screen

5.2.3 Results of Prediction (Presentation)

The Client Side consists of following pages.

Registration:

New row is inserted into user table of User database, after user submits the registration form.

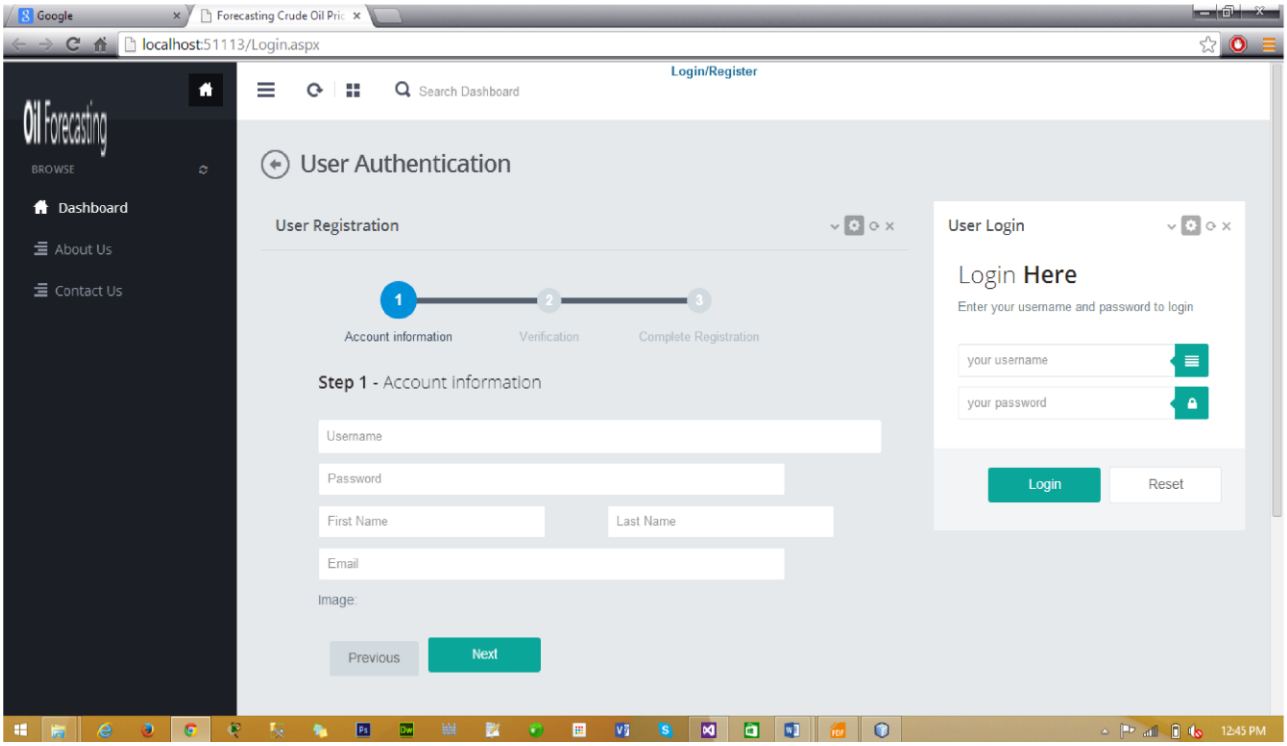

Figure 5-4 Registration Page

Login:

This feature allows the user to login. The user enters the username and password, these credentials are validated. If the username and password are correct, the user is logged in to the web application, else an error message is displayed and redirected to the Login page.

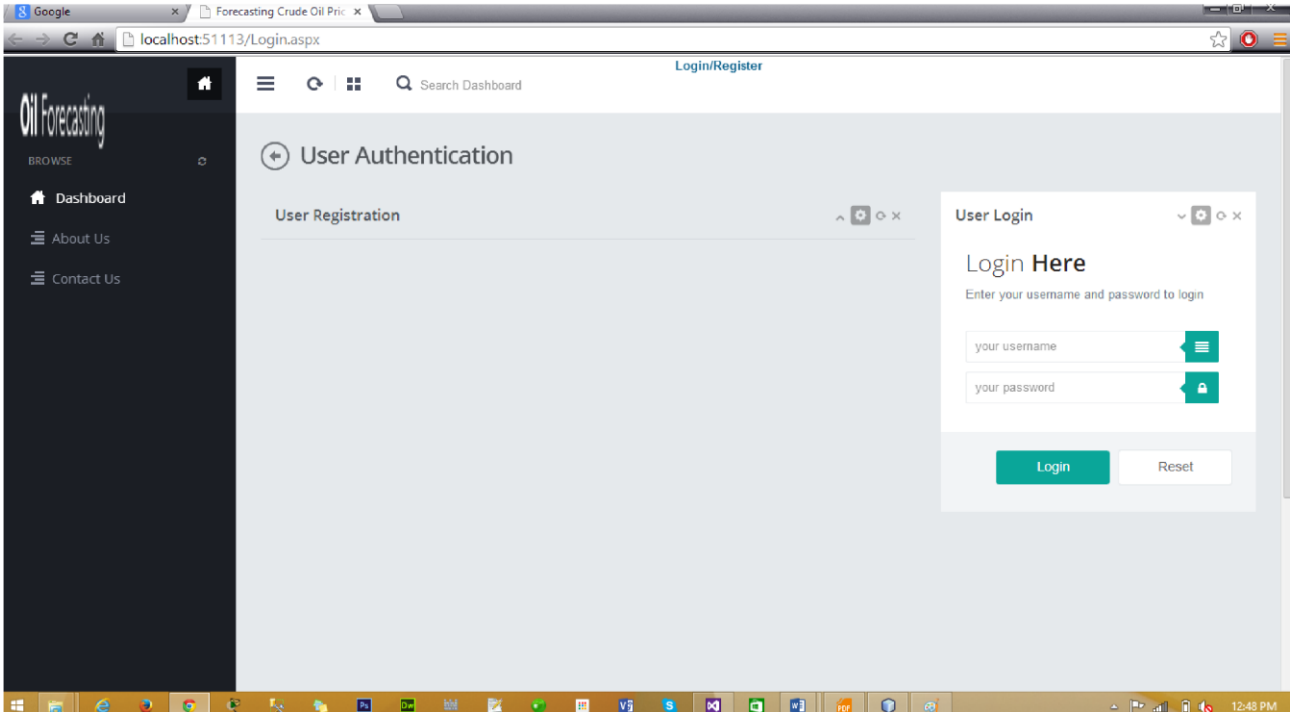

Figure 5-5 Login Page

Contact Us:

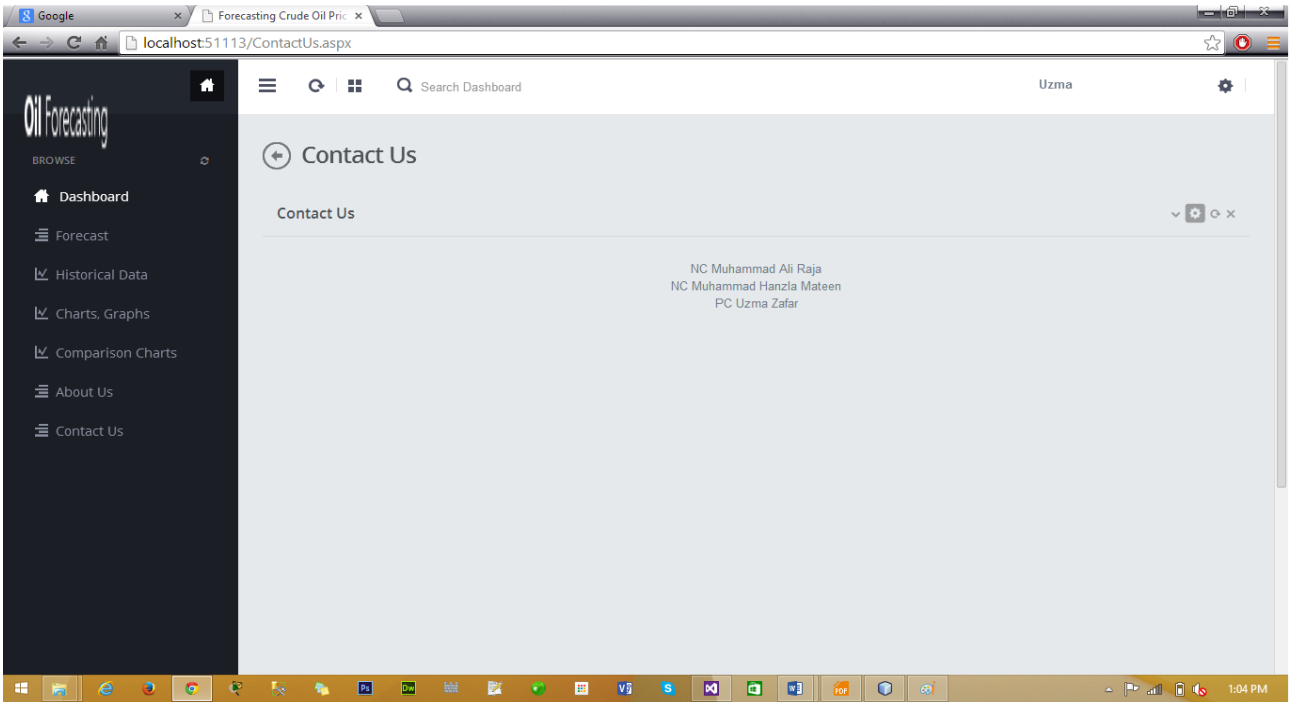

Figure 5-6 Contact Us Page

Dashboard:

This is the main screen of our client side. It displays information about the values of CPI, Supply, Stock and Production variables. It also displays the graph of oil price values of latest month.

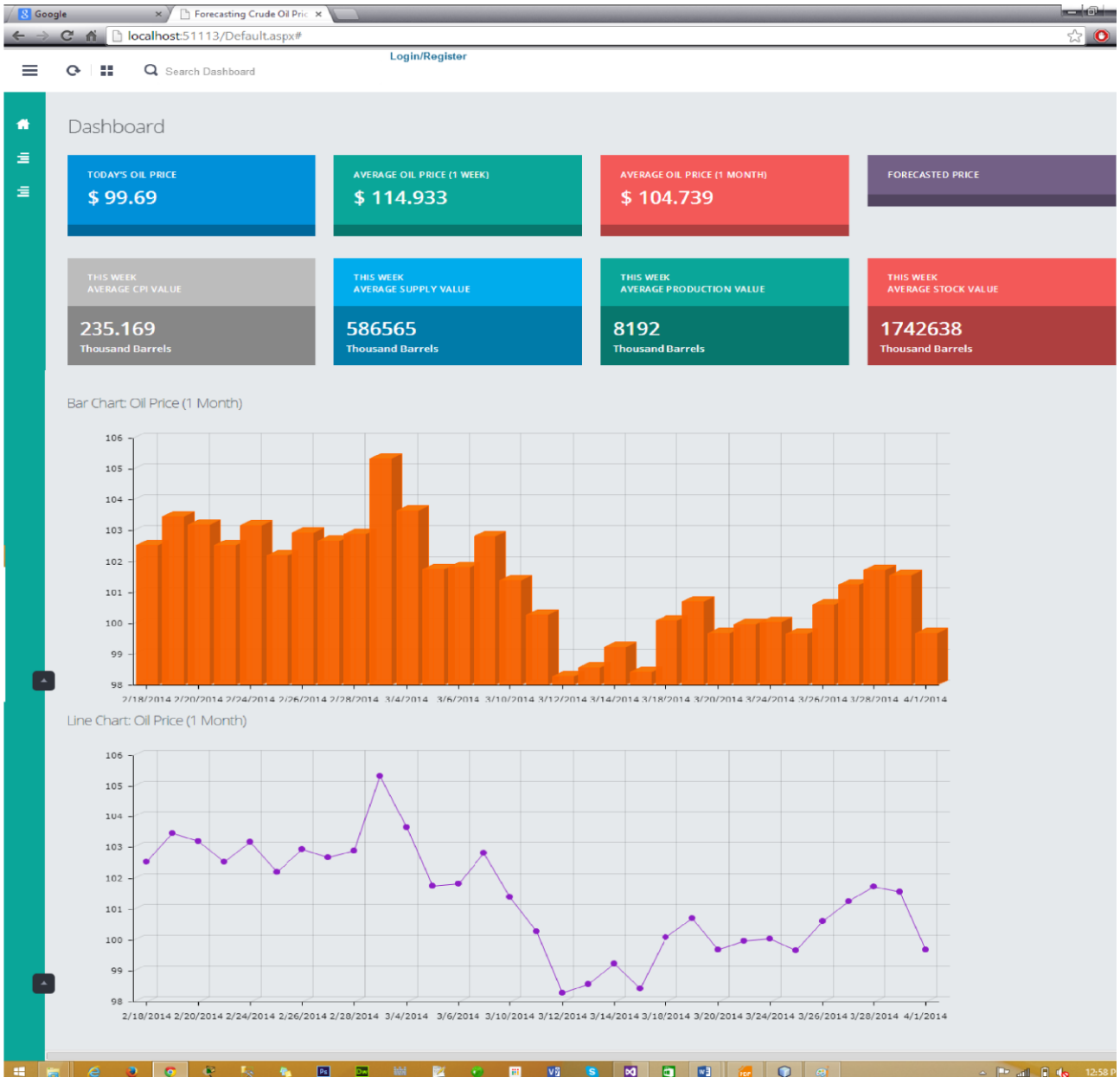

Figure 5-7 Dashboard

Generate Reports:

User can generate the reports. The user presses the Report button. The data is sent to the server side, and the data is fetched from data-vault database according to the details specified by the user. Then this data is sent to the client side and is displayed to the user.

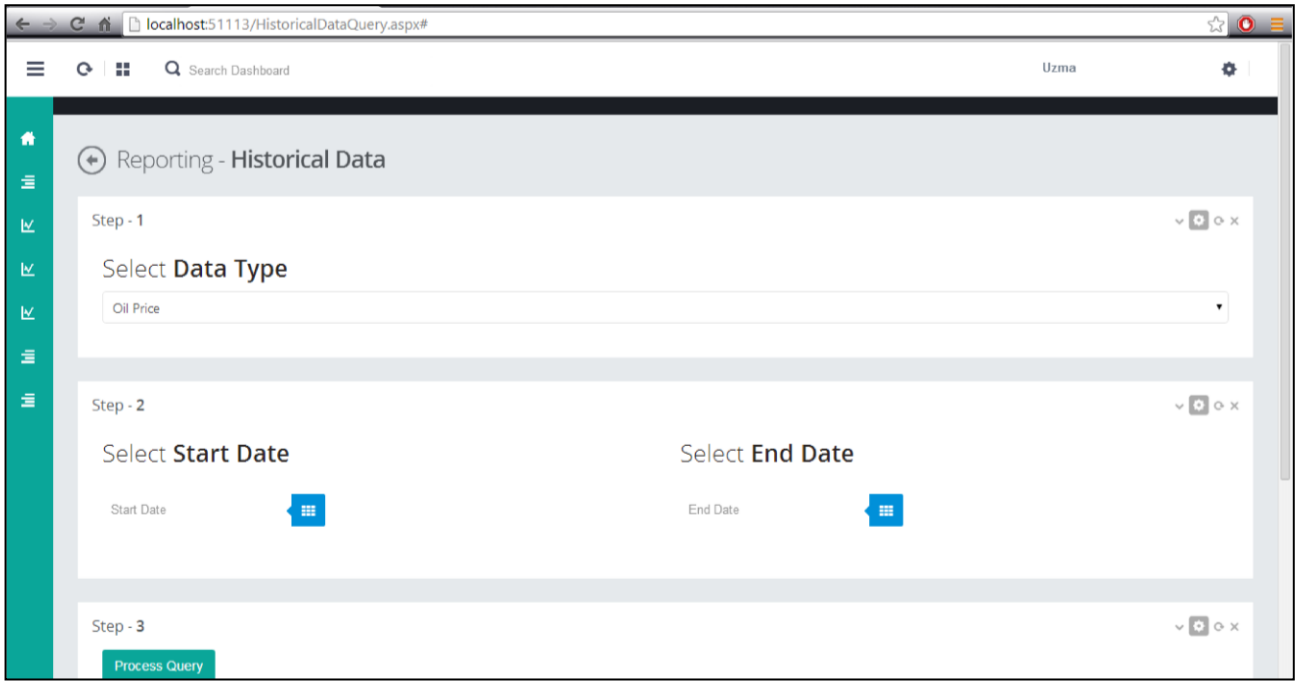

Figure 5-8 Generate Report Query Page

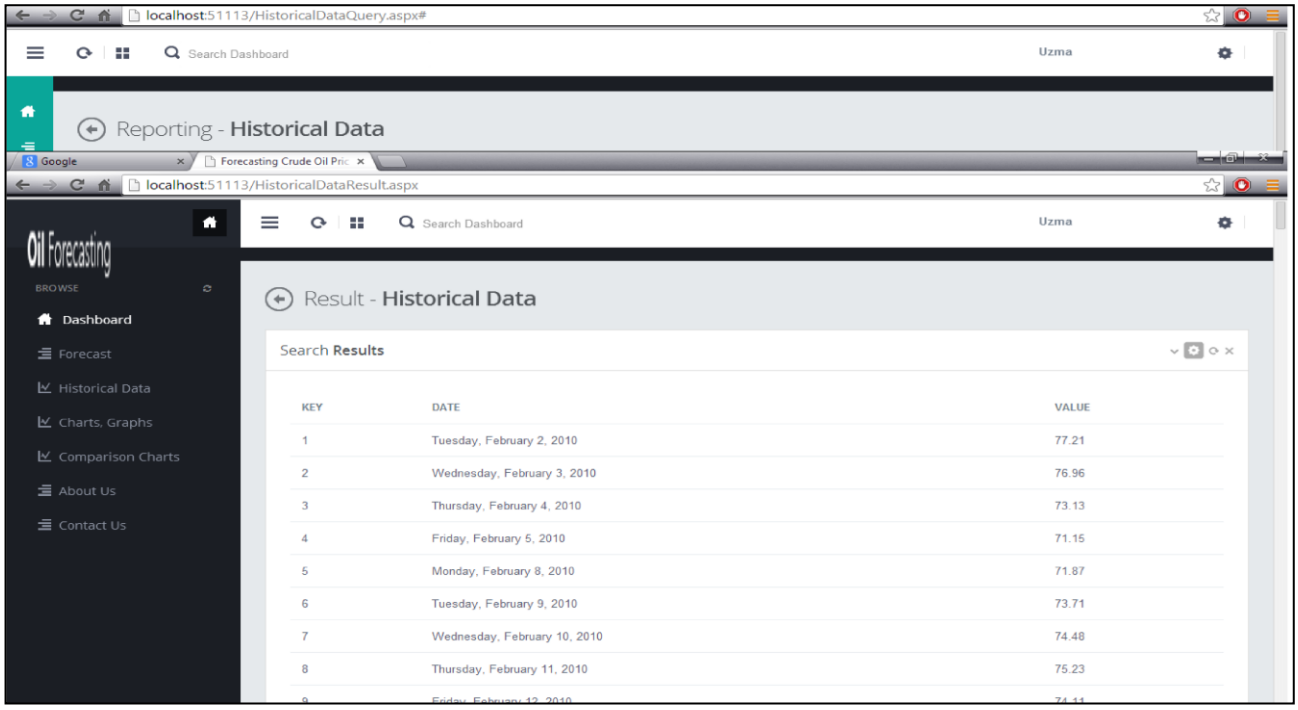

Figure 5-9 Generate Report Result Page

User can generate the graphs. The user presses the Graph button. The data is sent to the server side, and the data is fetched from data-vault database. Then this data is sent to the client side and is displayed to the user, in the form of the graph.

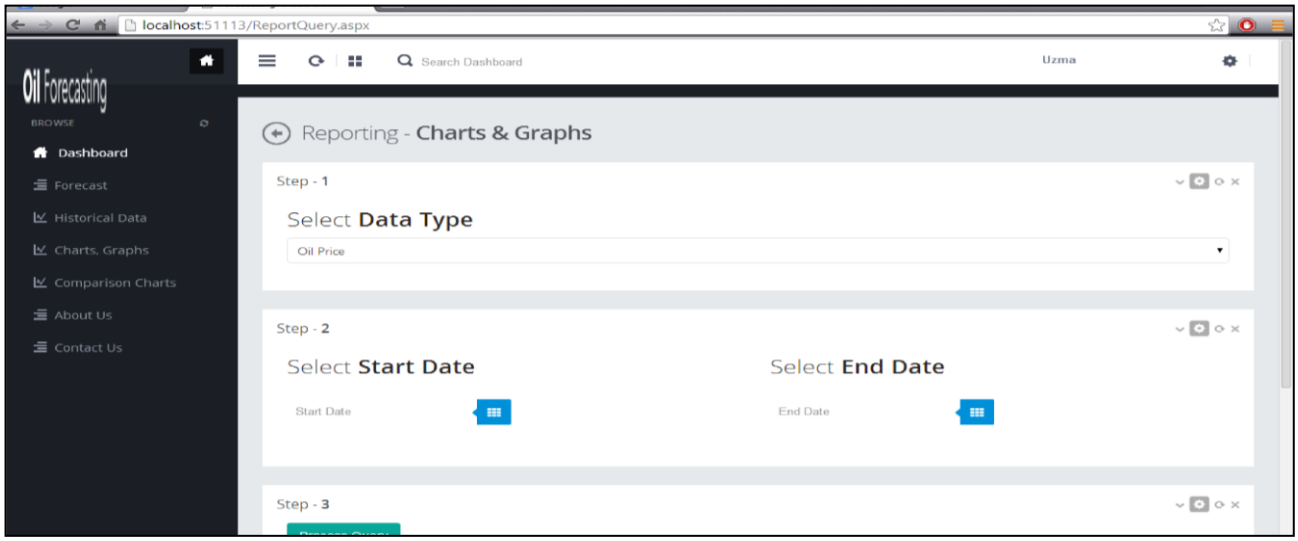

Figure 5-10 Generate Graph Query Page

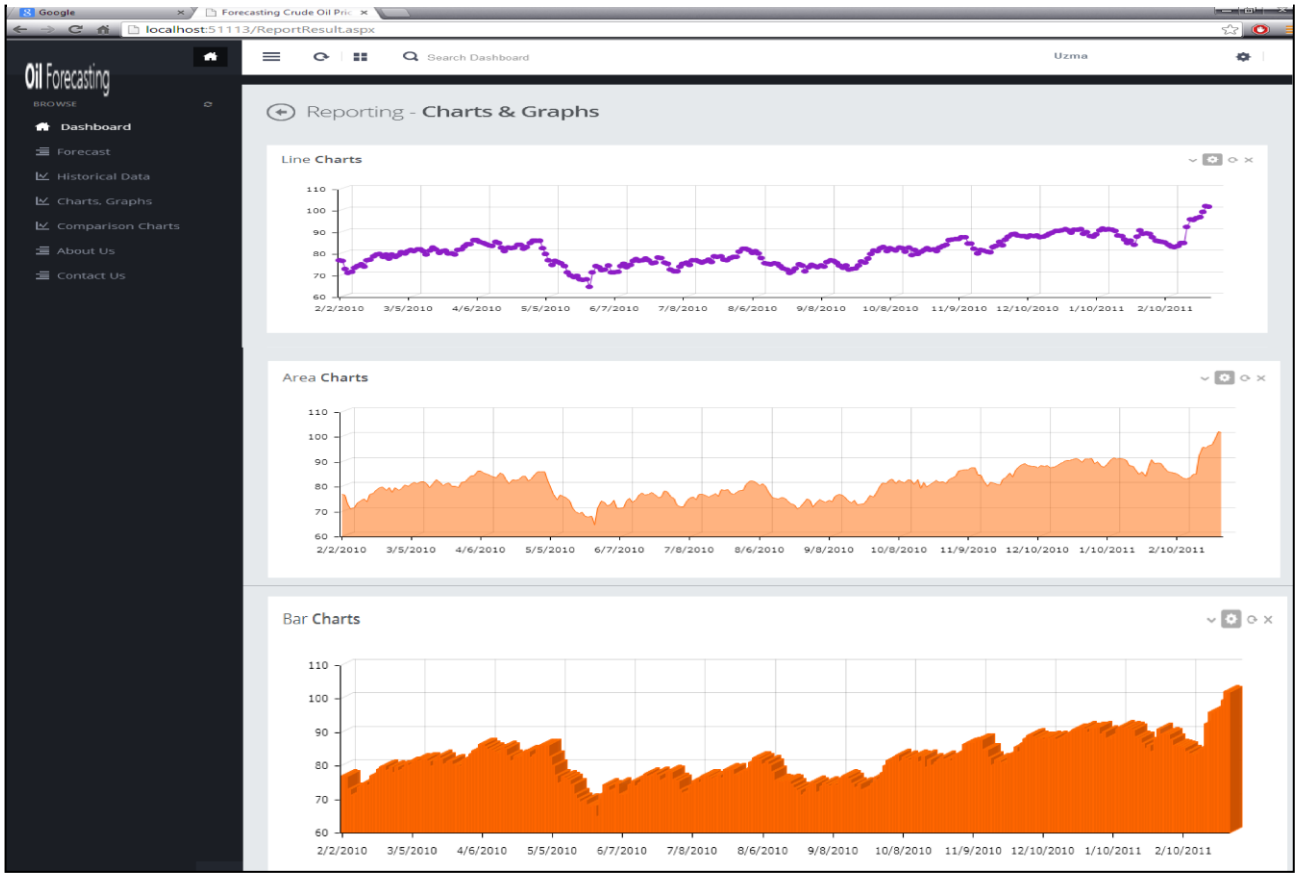

Figure 5-11 Generate Graph Result Page

Get Prediction:

User can see the prediction of Oil Prices. The user presses the Prediction button. The data is sent to the forecasting tool on the server side. The forecasting tool fetches data from data-vault database. The forecasting tool then forecasts the Oil Price on the specified date and sends this value to the client side and the value is displayed to the user.

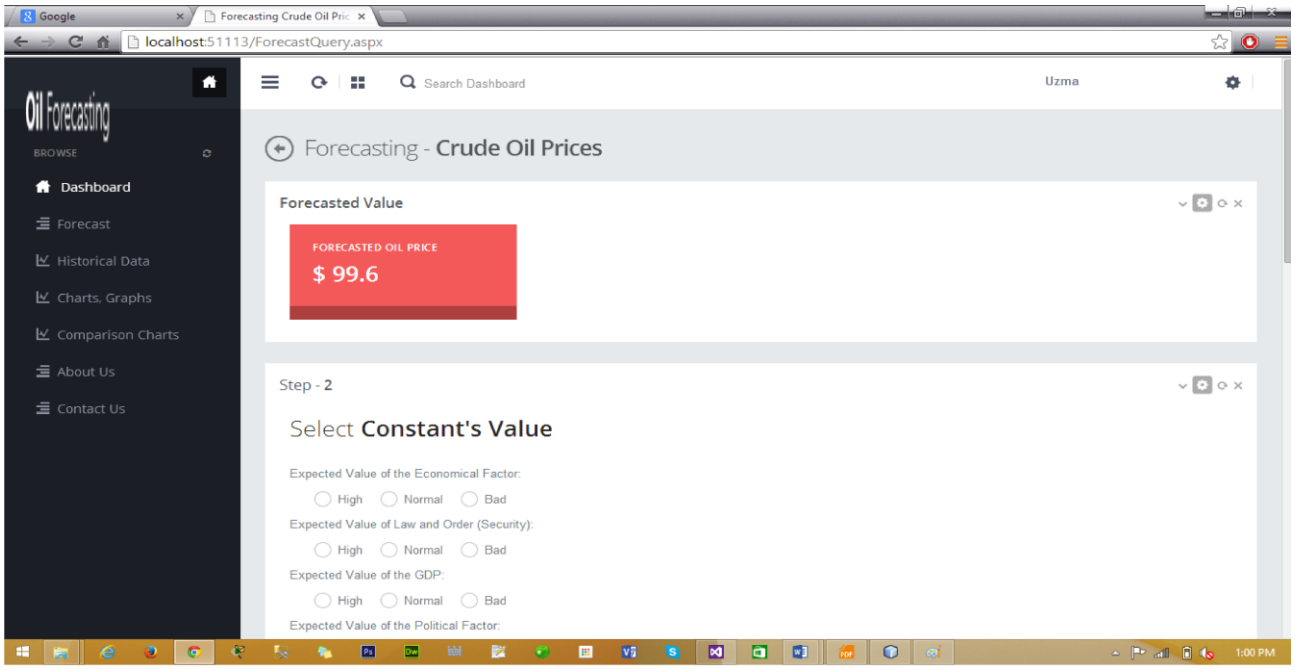

Figure 5-12 Forecasting Page

About Us:

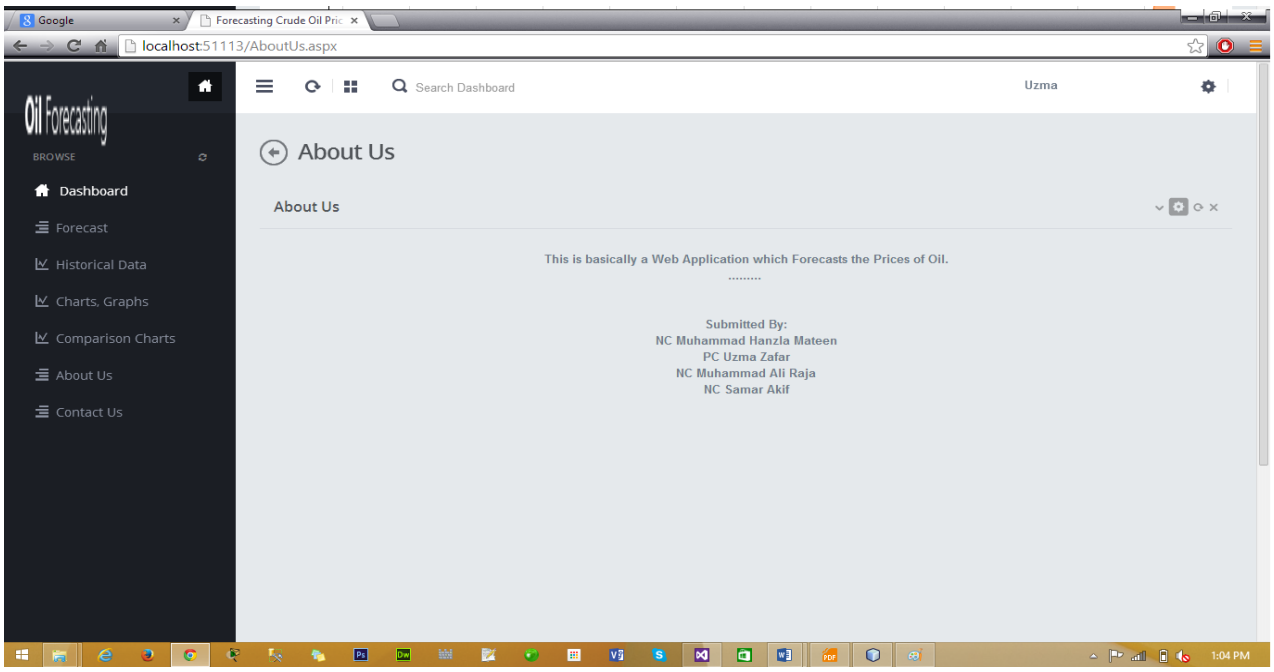

Figure 5-13 About Us Page

CHAPTER 6

TESTING AND RESULTS

6. Testing and Results

6.1 Overview

Testing of the software projects involve different levels of testing to make sure that the software which is being developed is error and fault free. Oil Price Prediction has different modules which were developed separately depending up on the functionalities. Therefore testing of all the modules has to be done and testing while integrating all the modules. The different levels at which testing was done are discussed here.

6.2 Unit Testing

Unit testing involves the testing of each module at the completion and sometimes during the development of the module. The test cases exercised as part of quality assurance of the application are:

6.2.1 Register User

Test Case Id	1
Unit to test	Register User
Assumptions	Clicked on Register Link User not previously registered User name available
Test Data	i. User Name ii. First Name iii. Last Name iv. Email ID v. Photo

Steps to be executed	Click on the register link Fill the registration form Click the next button
Expected Result	A confirmation message is displayed, and user is redirected to home page
Actual Result	A confirmation message is displayed, and user is redirected to home page
Pass/Fail	Pass

Table 6- 1 Register User Test Case

6.2.2 Login

Test Case Id	2
Unit to test	Login
Assumptions	User has not logged in previously Database holds user's record The login page has been opened
Test Data	i. User Name ii. Password
Steps to be executed	Open the login page Enter the user name Enter the password Click the login button
Expected Result	The user is redirected towards the homepage and the user name is displayed on the top left corner of the screen.
Actual Result	The user is redirected towards the homepage and the user name is displayed on the top left corner of the screen.
Pass/Fail	Pass

Table 6- 2 Login Test Case

6.2.3 Logout

Test Case Id	3
Unit to test	Logout
Assumptions	The user has already logged in to his account
Test Data	NULL
Steps to be executed	Click the logout link.
Expected Result	The user is redirected towards the homepage, registered user features are disabled
Actual Result	The user is redirected towards the homepage, registered user features are disabled
Pass/Fail	Pass

Table 6- 3 Logout Test Case

6.2.4 Forecasting Oil Price

Test Case Id	4
Unit to test	Forecast
Assumptions	User has logged in
Test Data	Fore-casted Value
Steps to be executed	Click on the forecast link, on the left side bar
Expected Result	The user is redirected towards the forecast page, on which the forecasted value of the oil price along with the relevant graphs is displayed.
Actual Result	The user is redirected towards the home page and the user name is displayed on the top left corner of the screen.
Pass/Fail	Pass

Table 6- 4 Forecasting Oil Price Test Case

6.2.5 Historical Data

Test Case Id	5
Unit to test	Historical Data
Assumptions	User has logged in Start date and end date are correct
Test Data	i. Type of data selected ii. Start Date iii. End Date
Steps to be executed	Click on the historical data link, on the left side bar Select type of data to be displayed (Oil price, CPI etc.) Select Start date Select End date Click Process Query button
Expected Result	The user is redirected towards the historical data page on which all the records of the particular data type are listed starting from the user defined start date and ending on the user defined end date.
Actual Result	The user is redirected towards the historical data page on which all the records of the particular data type are listed starting from the user defined start date and ending on the user defined end date.
Pass/Fail	Pass

Table 6- 5 Historical Data Test Case

6.2.6 Comparison Charts

Test Case Id	7
Unit to test	Comparison Charts
Assumptions	User has logged in Start date and end date are correct
Test Data	i. Type of data selected ii. Start Date iii. End Date
Steps to be executed	Click on the comparison chart link, on the left side bar Select type of data to be reported (Oil price, CPI etc.) Select Start date Select End date Click Process Query button
Expected Result	The user is redirected towards the comparison charts page, on which numerous graphs show the actual value against the forecasted value.
Actual Result	The user is redirected towards the comparison charts page, on which numerous graphs show the actual value against the forecasted value.
Pass/Fail	Pass

Table 6- 6 Comparison Charts Test Case

6.2.7 Extract Transform Load Module (Files Download)

Test Case Id	8
Unit to test	Files(ETL TOOL)
Assumptions	It is assumed that the save locations have already been specified, as well as the download locations
Test Data	Type of variable selected.

Steps to be executed	In the ETL TOOL click on the files tab Click the checkbox of the variable you wish to download excel data file. Click on the process button.
Expected Result	A success notification is shown.
Actual Result	A success notification is shown.
Pass/Fail	Pass

Table 6- 7 Extract Transform Load Module (Files Download) Test Case

6.2.8 Extract Transform Load Module (Extract)

Test Case Id	9
Unit to test	Extract(ETL TOOL)
Assumptions	It is assumed that the data source path has already been specified, and data of all the variables exists at that path.
Test Data	Type of variable selected.
Steps to be executed	In the ETL TOOL click on the extract tab Click the checkbox of the variable you wish to extract Click on the process button.
Expected Result	A success notification is shown.
Actual Result	A success notification is shown.
Pass/Fail	Pass

Table 6- 8 Extract Transform Load Module (Extract) Test Case

6.2.9 Extract Transform Load Module (Load)

Test Case Id	10
Unit to test	Load(ETL TOOL)
Assumptions	It is assumed that the data source path has already been specified, and data of all the variables exists at that path.
Test Data	Type of variable selected.
Steps to be executed	In the ETL TOOL click on the load tab Click the checkbox of the variable you wish to extract Click on the process button.
Expected Result	The processes going on at the back end are displayed in a list below, and success message is displayed.
Actual Result	The processes going on at the back end are displayed in a list below, and success message is displayed.
Pass/Fail	Pass

Table 6- 9 Extract Transform Load Module (Load) Test Case

6.2.10 Forecasting Module

Test Case Id	11
Unit to test	Forecasting Tool
Assumptions	It is assumed that the data of the variables on which forecast is performed is up to date.
Test Data	Specified date range's value for each input variable. No of hidden layers and nodes
Steps to be executed	In the FORECASTING TOOL after specifying the inputs click train to perform the training.
Expected Result	If the training is performed successfully log will show success.
Actual Result	Training is performed successfully log showed success.
Pass/Fail	Pass

Table 6- 10 Forecasting Tool Test Case

6.3 Integration Testing

6.3.1 Increment 1: Registration Module

Test Case ID	Input(s)	Initial Condition	Output(s)	Sequence of Actions	Result
12	i. User Name ii. First Name iii. Last Name iv. Email ID v. Photo	System is running and database is connected	User registered	i. Click on the register link ii. Fill the registration form iii. Click the next button	PASS

Table 6- 11 Integration Testing Registration Module

6.3.2 Increment 2: Integration with User Login Module

Test Case ID	Input(s)	Initial Condition	Output(s)	Sequence of Actions	Result
13	i. User Name ii. First Name iii. Last Name iv. Email ID v. Photo	System is running and database is connected	User registered Login Page Appears	i. Click on the register link ii. Fill the registration form iii. Click the next button iv. Account Created	PASS
14	i. Username **ii.** Password	System is running and database is connected	System user status Logged in Dashboard Page appears	i. Open the login page ii. Enter the user name iii. Enter the password iv. Click the login button	PASS

Table 6- 12 integration Testing User Login Module

6.3.3 Increment 3: Integration with Forecasting Module

Test Case ID	Input(s)	Initial Condition	Output(s)	Sequence of Actions	Result
15	**i.** Username ii. Password	System is running and database is connected, and user logged in	User logged in	i. Open the login page ii. Enter the user name iii. Enter the password iv. Click the login button	PASS
16	**i.** Username **ii.** Password	System is running, database is connected and user is logged in	Forecasted value is shown	Click on the forecast link, on the left side bar	PASS

Table 6- 13 Integration Testing Forecasting Module

6.3.4 Increment 3: Integration with Data Entry Module

Test Case ID	Input(s)	Initial Condition	Output(s)	Sequence of Actions	Result
17	i. Start Date ii. End Date iii. Type iv. Process query button click	System is running and database is connected and user is logged in	Next page appears after selected option from left side bar	i. Enter Start Date ii. Enter End Date iii. Enter Type	PASS

Table 6- 14 Integration Testing Data Entry Module

6.3.5 Increment 4: Integration with Historical Data module

Test Case ID	Input(s)	Initial Condition	Output(s)	Sequence of Actions	Result
18	i. Start Date ii. End Date iii. Type iv. Process query button click	i. System is running, database is connected ii. User is logged in	Next page appears after selected option from left side bar	i. Enter Start Date ii. Enter End Date iii. Enter Type	PASS
19	No input	System is running, database is connected, and user has logged in, and Input submitted.	Historical data is displayed in list form	i. Enter Start Date ii. Enter End Date iii. Enter Type iv. Proceed button	PASS

Table 6- 15 Integration Testing Historical Data module

6.3.6 Increment 5: Integration with Charts module

Test Case ID	Input(s)	Initial Condition	Output(s)	Sequence of Actions	Result
20	i. Start Date ii. End Date iii. Type iv. Process query button click	i. System is running, database is connected ii. User is logged in	Next page appears after selected option from left side bar	i. Enter Start Date ii. Enter End Date iii. Enter Type	PASS
21	No input	System is running, database is connected, and user has logged in.	Charts are displayed	i. Enter Start Date ii. Enter End Date iii. Enter Type iv. Proceed button	PASS

Table 6- 16 Integration Testing Charts Module

6.3.7 Increment 6: Integration with Comparison Charts module

Test Case ID	Input(s)	Initial Condition	Output(s)	Sequence of Actions	Result
22	i. Date ii. Process query button click	i. System is running, database is connected ii. User is logged in	Next page appears after selected option from left side bar	Enter Date	PASS
23	No input	System is running, database is connected, and user has logged in, and Input submitted.	Comparison Charts are displayed	i. Enter Date ii. Proceed button	PASS

6.4 System Testing

System testing is the level of testing which comes when the whole system has been developed and integrated. The complete system was tested in different inputs with different conditions to verify that those conditions do not disrupt the performance of the system. There were no significant findings that need to be documented.

CHAPTER 7
CONCLUSION

7. Conclusion

Existing solutions rely heavily on the use of neural networks. Accuracy to some extent may be achieved by the existing systems, however they lack in the amount of detailed information that a user requires in order to make intelligent business decisions. Data is simply displayed in a list form and there is little use of graphical representation.

The systems that are currently available have following shortcomings:

i. Only bimonthly reports are generated.
ii. Graphs are not utilized efficiently.
iii. Day to day analysis is not possible.

The above mentioned ways are therefore not accurate as well as specific. There is a need to find a solution that can give more detailed information, proper usage of graphs to exhibit our analysis in a proper way. Also the system should provide reporting considering both managerial users as well as other users.

Proposed solution combines the concepts of Data Warehousing, Time Series and Neural Networks and uses graphs and reporting effectively. It is a Web Application, the user will communicate with the system through web pages using web browser, and there is also a server-side component which will be responsible for database and synchronization services. The scope of the project encompasses both server and client-side functionalities. Data is gathered from resources and a data warehouse is created where all the historic data is stored.

Our System gives 97% accurate results.

In short through this project we have learned to apply all software engineering steps learned during the course on practical application. This project has also introduced us to web and its technologies. It has also taught us to work on a project as team and to coordinate with team members.

8. Future Work

Attributes like dollar prices, gold prices which make a great effect on oil market upward and downward trends can also be introduced as external variables or input from user.
The following factors that are being considered constant, can be made variable.

i. Value of the Economical Factor
ii. Value of Law and Order (Security)
iii. Value of the Political Factor

News and articles are best source to get political situation. Text mining field of data mining can be incorporated, which is the extraction of relevant materials from news and articles about political issues and other variables.
Sending customer about news and predictions as email/messages so user is kept informed of current trends and future predictions.

9. Bibliography

i. http://en.wikipedia.org/wiki/Time_series *(accessed on 7-October-2013)*

ii. http://en.wikipedia.org/wiki/Data_warehouse *(accessed on 7-October-2013)*

iii. http://www.statsoft.com/textbook/time-series-analysis/

iv. http://www.tutor2u.net/economics/revision-notes/as-markets-oil.html

v. https://www.imf.org/external/np/speeches/2009/031809.htm

vi. http://www.ask.com/question/factors-affecting-oil-prices

vii. http://www.economicshelp.org/blog/3809/oil/factors-affecting-oil-prices-in-short-term-and-long-term/

viii. http://www.investmentu.com/2012/January/the-five-factors-moving-oil-prices-this-year.html

ix. http://c1wsolutions.wordpress.com/2012/04/30/factors-affect-price-of-oil/

x. http://www.fool.com/investing/general/2013/04/21/3-factors-that-dramatically-affect-the-price-of-oi.aspx

xi. www.ijcte.org/papers/504-G1345.pdf

xii. www.psrcentre.org/images/extraimages/413696.pdf

xiii. www.stat.tamu.edu/~fliang/BNNtime.pdf

xiv. www.arxiv.org/pdf/0906.4838

xv. www.igi-global.com/chapter/data-mining-methods-crude-oil/44289

xvi. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5596602&tag=1

xvii. www.ncbi.nlm.nih.gov/pubmed/18252536

xviii. www.cesifo-group.de/DocCIDL/forum1-09-focus3.pdf

xix. www.oenb.at/en/img/mop_20042_the_impact_tcm16-19678.pdf

xx. www.ijcte.org/papers/504-G1345.pdf

xxi. www.psrcentre.org/images/extraimages/413696.pdf

xxii. www.stat.tamu.edu/~fliang/BNNtime.pdf

xxiii. www.arxiv.org/pdf/0906.4838

xxiv. www.igi-global.com/chapter/data-mining-methods-crude-oil/44289

xxv. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5596602&tag=1

xxvi. www.ncbi.nlm.nih.gov/pubmed/18252536

xxvii. www.cesifo-group.de/DocCIDL/forum1-09-focus3.pdf

xxviii. www.oenb.at/en/img/mop_20042_the_impact_tcm16-19678.pdf

xxix. http://www.hindawi.com/journals/jam/2012/646475/

xxx. https://www.google.com.pk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&ved=0CDMQFjAB&url=http%3A%2F%2Fwww.sysmath.com%2Fjweb_xtkxyfzx%2FEN%2Farticle%2FdownloadArticleFile.do%3FattachType%3DPDF%26id%3D11363&ei=XrH1Ua28M-3P4QSjuYCYBA&usg=AFQjCNENbaJeF1eB7rUn05_2vLwQgJPeWw&sig2=ONNzgvKv1Dr1BnKMJHZ3iA

xxxi. http://www.ijcte.org/papers/504-G1345.pdf

xxxii. https://www.google.com.pk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&ved=0CDMQFjAB&url=http%3A%2F%2Fwww.sysmath.com%2Fjweb_xtkxyfzx%2FEN%2Farticle%2FdownloadArticleFile.do%3FattachType%3DPDF%26id%3D11363&ei=XrH1Ua28M-3P4QSjuYCYBA&usg=AFQjCNENbaJeF1eB7rUn05_2vLwQgJPeWw&sig2=ONNzgvKv1Dr1BnKMJHZ3iA

xxxiii. WANG Shouyang (Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing100080,China; School of Management, Graduate School of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100049,China. Email: swang@iss.ac.cn)

xxxiv. S. N. Abdullah, X. Zeng, School of Computer Science. The University of Manchester, Oxford Road, Manchester, M13 9PL United Kingdom

xxxv. FAMING LIANG Department of Statistics, Texas A&M University, College Station, TX 77843- 3143, USA fliang@stat.tamu.edu Received April 2002 and accepted May 2004

xxxvi. GILIOLAFREY* MATTEOMANERA* ANILMARKANDYA* ELISASCARPA****

xxxvii. Amir F. Atiya, Senior Member, IEEE, Suzan M. El-Shoura, Samir I. Shaheen, Member, IEEE, and Mohamed S. El-Sherif