Mudiaga Innocent Oke

Comparative Analysis To Determine Predictive Model Accuracy

A dynamic currency exchange rate predictive model development using SAP HANA Predictive Analytic Library (PAL) algorithm

Helsinki Metropolia University of Applied Sciences

Masters of Business Administration

Business informatics

Master's thesis

10.03.2014

Helsinki Metropolia University of Applied Sciences

Author(s)	Mudiaga Innocent Oke
Title	Comparative Analysis To Determine Predictive Model Accuracy
Number of Pages	80 pages + 17 appendices
Date	10.03.2014
Degree	Master of Business Administration
Degree Programme	Degree Programme in Business Informatics
Specialisation option	
Instructor(s)	Antti Hovi, Senior Lecturer

The present thesis describes the development and implementation of a dynamic currency exchange rate predictive model. The aim of the thesis was to measure and determine the accuracy of a dynamic currency exchange rate predictive model by analysing different historical data samples.

The theoretical framework of the thesis focused on research into different disciplines related to predicted analytics and the different data mining algorithms. The study was carried out using quantitative data samples and SAP high performance analytic appliance predictive analysis library (PAL) Time series double exponential algorithm. The measurement was done by comparing the predicted or forecasted exchange rates against the actual exchange rates. Standard statistical methods were used to determine the accuracy of the predictive model.

The results of the study showed that last three months data sample or most recent data gives better predictive results for short term forecasting while the full data sample or entire data set gives better result for longer term forecasting.

Based on the study, it is recommended that fundamental analysis of currency exchange method which takes account of the driving forces behind currency exchange rates such as political and economic situation, the rise and fall of interest rates and other economic indicators should be incorporated along technical analysis which involves the use of historical data to get give better accuracy.

Keywords	Predictive analytics, SAP HANA, Double exponential algorithms, Data mining, Model, forecasting, Business Intelligence

Contents

Appendices

List of figures

List of tables

Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network
BA	Business Analytics
BI	Business Intelligence
BW	Business Warehouse
DES	Double Exponential Smoothing
DM	Data Mining
DW	Data Warehouse
GA	Genetic Algorithm
HANA	High Performance Analytics Appliance
IT	Information Technology
KDD	Knowledge Discovery from Data
MAE	Mean Absolute Error
OLAP	Online Analytic Processing
OLTP	Online Transaction Processing
PA	Predictive Analytics
PAL	Predictive Analytics Library
RAM	Random Access Memory
RDBMS	Relational Database Management System
RMSE	Root Mean Square Error
RSE	Relative Square Error
SAP	Systems Application Programming
SES	Single Exponential Smoothing

1 Introduction

This chapter discusses and provides insights to background of this master's thesis, research purpose and question and finally it covers the overall structure of the thesis.

1.1 Background

Technology for business intelligence has changed. The breakthrough of new technologies such as In-Memory computing database and innovation of advance software have enabled businesses to store, process and analyse a massive amount of data in real-time as never before seen. The way information is collected, stored, processed and analysed has never been this efficient with the help of these new technologies and advance software applications.

In today's business environment, for companies to be competitive in business environment, adapt to changes in the market and stay ahead of their competitors, it is simply not enough to make decisions based on past and present information which only provides information to what has happened using traditional analytics applications or rely on decision making based on the gut feeling of senior management (Khan et al., 2008: 581).

Zaman (2013:1) stated that business organizations need to know more about the future and in particular, future trends, patterns and customer behavior in order to understand the business climate and the market in which they operate better. Figure 1 depicts that with predictive analytics, a business should be able to predict what will or might happen and make informed decisions unlike with traditional reporting, analytics and monitoring tools.

Business intelligence disciplines such as data mining and predictive analytics which can predict what will or might happen in the future, as shown in figure 1, are used to extract both structured and unstructured data from different heterogeneous sources and uncover hidden patterns and relationship of similar information.

Predictive analytics is part of business intelligence which combines techniques from statistics, modelling, machine learning and data mining that uses historical data and

operational data to forecast or make predictions about future occurrences. Predictive analytics is referred to as a model ability to generate accurate predictions of new observations, where new can be interpreted temporally as observations in a future time period (Shmueli et al, 2011:9). Gaultieril (2013:1) said that predictive analytics uses algorithms to find patterns in data that might predict similar outcomes in the future.

The core element of predictive analytics is the predictor which is a variable that can be measured for an individual or entity to predict future behaviour. For example, a credit card company could consider age, income, credit history, and other demographics as predictors when issuing a credit card to determine an applicant's risk factor (Zaman, 2013:1).

Companies such as Yahoo, Facebook, LinkedIn, Google and Netflix have been collecting and analysing data about their users and customers. The data collected are used to improve the quality of services provided for their users and retaining their customers. The data collected by the companies from ubiquitous sources such as the geo-location tagging in smartphones, information posted on social networking sites, online chatting and blogging by users can be a potential goldmine if harnessed and utilized properly. They have realised that there are lots of opportunities to be tapped with the information collected because they can increase and get better insight about their customers, competitors and business and generate new business opportunities.

Common practical applications of predictive analytics include the use of big data and predictive analytics by LAPD police to prevent a crime before it happens. Another notable application of predictive analytics is credit scoring by financial services, weather predictions by meteorological institutes and customers churning by telecommunication companies.

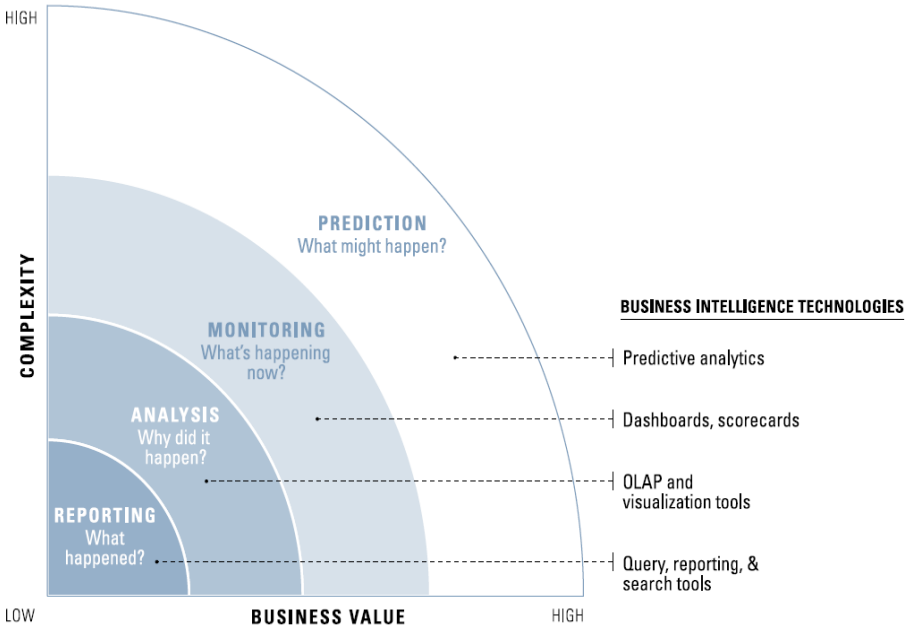

Figure 1 – The spectrum of BI Technologies (Wayne W. Eckerson, 2007:5)

1.1 Research purpose and motivation

The present thesis work is purely for academic research and for the author's professional growth. The author will use this thesis work and knowledge gained during the practical applications and implementations derived from this thesis research to get better insights and understanding of next generation of business intelligence technologies. Gartner (2013:1) stated in figure 2 that the plateau of productivity of predictive analytics will become mainstream and will be reached within next 2 years. What this means is that predictive analytics adoption will become a common practice by companies to enable them to make better choices and decisions.

In making predictions about future events or occurrences, there is hardly a correct answer as the purpose is to make a good guess as much as possible and be less wrong instead of getting completely wrong result. Random walk theory has also stated that it is impossible to make predictions with historical data for example predicting stock prices or currency exchange rates.

The purpose of this thesis is to compare and analyze different data samples of historical and current data to determine which data sample gives better accuracy of a predictive model. In order to achieve this, the goal will be to develop currency exchange forecasting model using quantitative data sets and SAP High performance analytics appliance (HANA) predictive analytic library software and compare the predicted or forecasted result against the actual result. Other ways of improving the model accuracy will also be discussed.

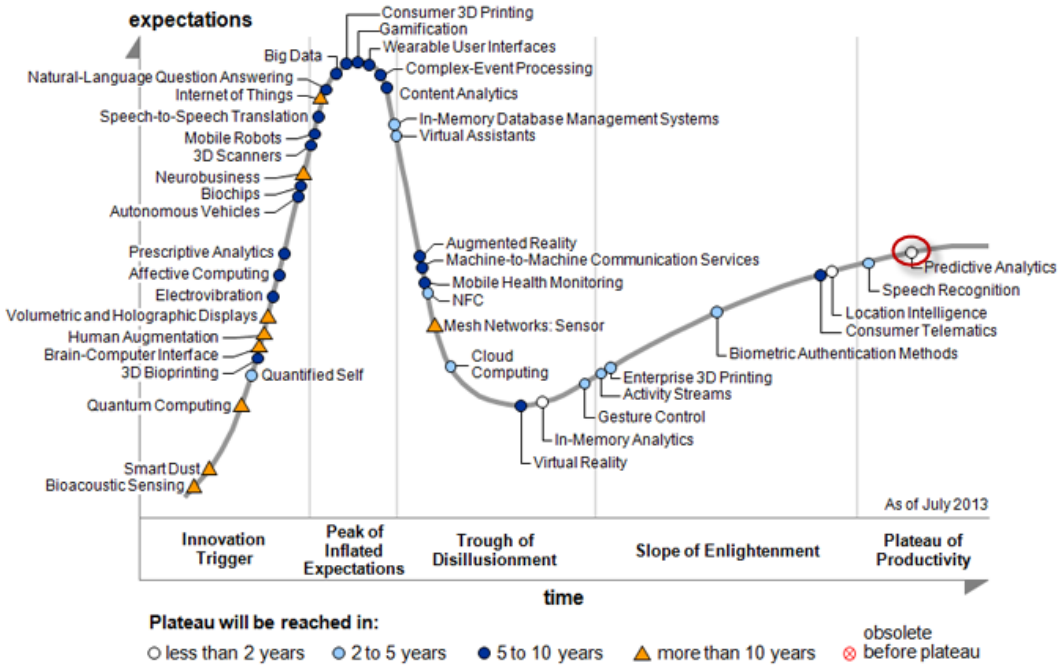

Figure 2 - Gartner Hype Cycle for Emerging Technologies, 2013

1.2 Research question

The research question is how to determine the accuracy of a predictive model. The case will be based on currency forecasting model.

1.3 Thesis Structure

Chapter 1 begins with a general background about predictive analytics and the importance of the discipline. The research purpose and question are also discussed.
Chapter 2 begins with the discussion of the different disciplines associated with predictive analytics such as business intelligence, data warehouse, in-memory computing, and data mining and predictive analytics.
The different predictive analytics algorithms will be introduced in chapter 3. However, the mathematics behind the different algorithms and how they aid in making predictions will not be discussed in detail.
The evaluation of predictive model performance will be discussed in chapter 4 and chapter 5 will cover the predictive model development using SAP HANA PAL. Finally, conclusions and recommendations will be discussed in chapter 6.

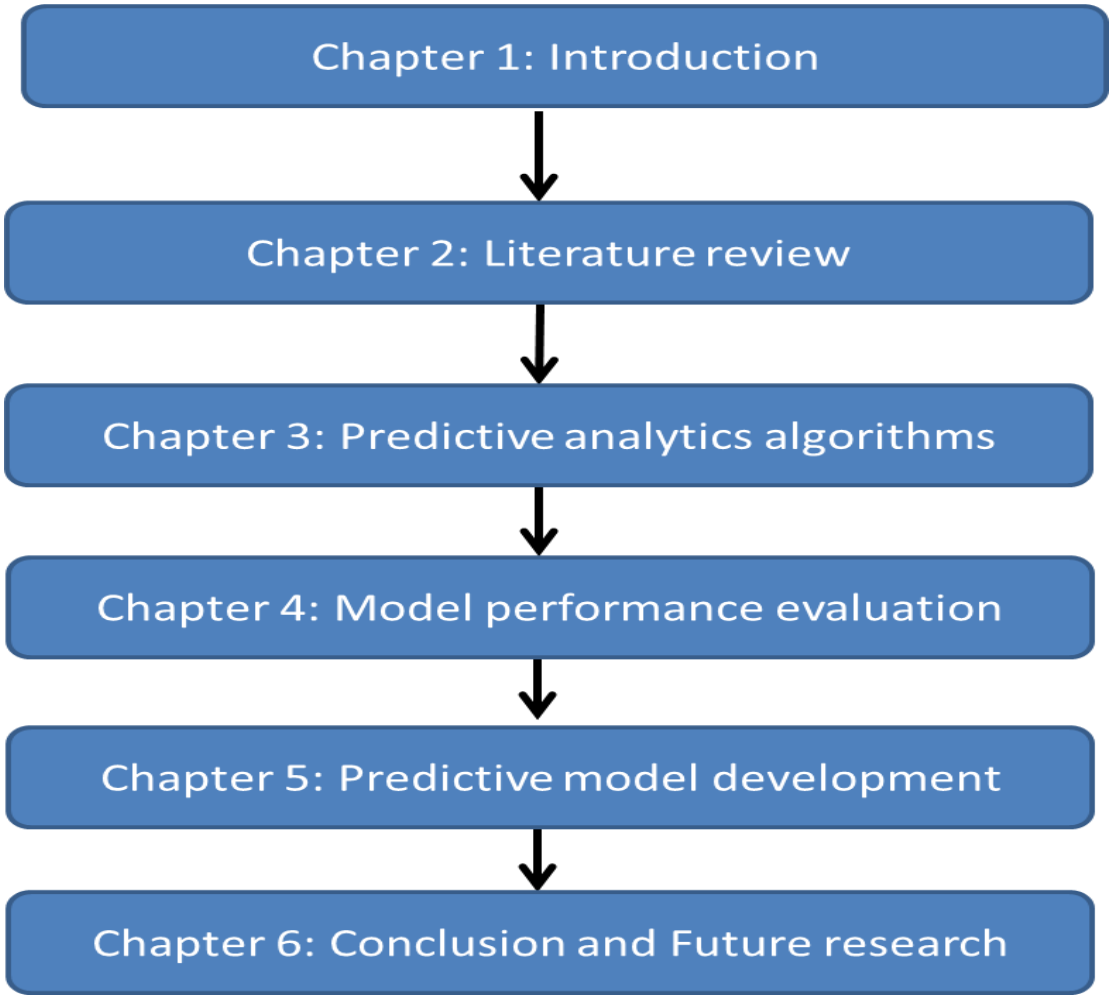

Figure 3 - Thesis structure

2 Literature review

This chapter discusses the concepts of machine learning, big data, data warehouse, business intelligence and in-memory computing and other disciplines associated with predictive analytics.

2.1 Machine Learning

A recent report from the Mckinsey Global institute asserts that machine learning a.k.a data mining or predictive analytics will be the driver of the next big wave of innovation ((Pedro Domingos,2013:1). Machine learning is a branch of business intelligence concerned with the design and development of algorithms that allows computers to evolve behaviours based on empirical data (James Manyinka et. al; 2011:29).

One of the main purpose of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data and some of the successful applications includes speech recognition, bio-surveillance, computer vision and robot control (Tom Mitchel, 2006:4).

2.2 Big Data

Intel IT Center (2013) refers big to huge datasets that are orders of magnitude larger (volume); more diverse, including structured, semi-structured and unstructured data (variety) and arriving faster (velocity) than what organizations or individuals have dealt with before.

The capturing, storing, managing and analysing of big data is beyond the ability of typical database software tools (Jame Manyika et al.; 2011:1).However, with in-memory computing technologies big data can easily be analysed and processed to generate real time information.

Big data is characterized by 3Vs as seen in the figure below.

- Volume – The massive scale and growth of unstructured data outstrips traditional storage and analytical solutions.
- Variety – Traditional data management processes can't cope with the heterogeneity of big data
- Velocity – data is generated real time with demands of usable information to be served immediately.

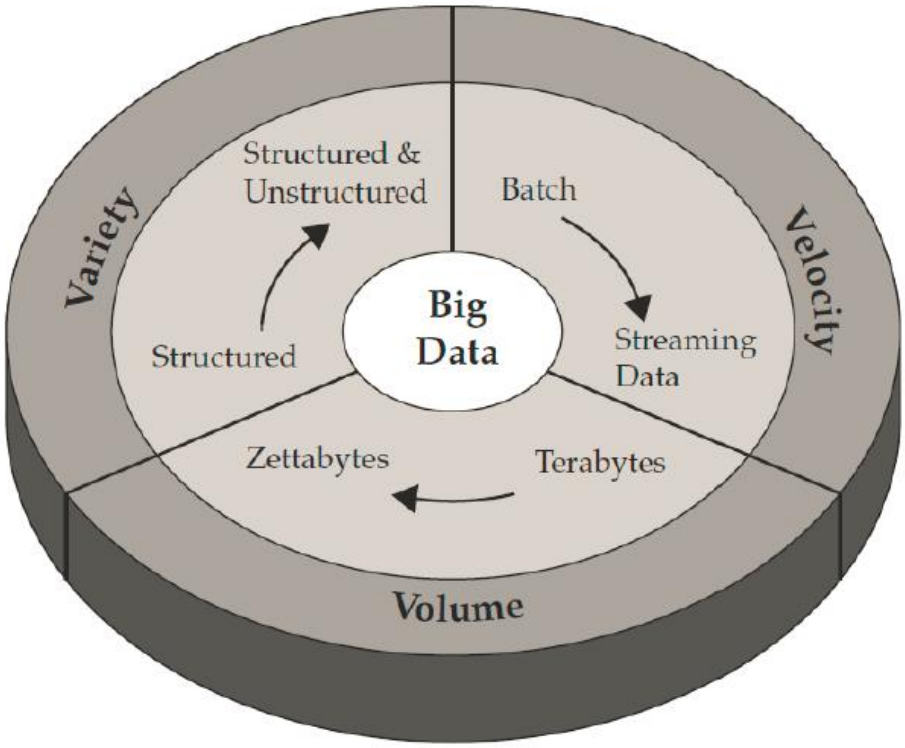

Figure 4 - 3Vs characterization of big data (Marko Grobelnik, 2012)

The flood of data is generated by connected devices from PCs and smartphones to sensors such as RFID readers and traffic cams. Plus it is heterogeneous and comes in many different formats including text, documents, images, videos and more (Intel IT Center, 2013:3).

It is predicted that by 2015, there will be about 8ZB (Zeta bytes) of data as shown in below figure and nearly 15 billion connected devices (ibid).

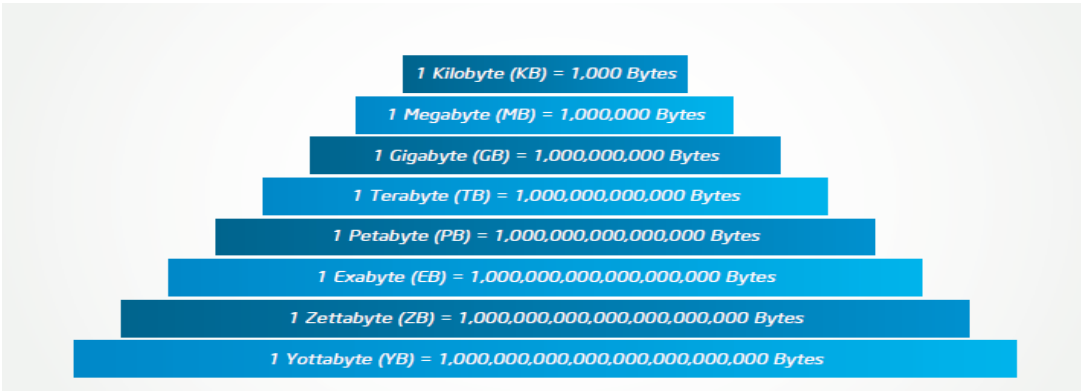

Figure 5 - Data size of big data (Intel IT center, 2013)

2.3 Data Warehouse

The term data warehouse was first coined in the early 1990s by Bill Inmon who defined data warehouse as “a subject-oriented, integrated, time variant, non-volatile collection of data organized to support management needs” (Zhenyu et al. 2002:23).

According to (ibid.), data warehousing assists organizational information processing by providing a solid platform of integrated data both current and historical, from which organizations can conduct a series of business analyses.

Data transferred to data warehouse usually come from various online transaction processing (OLTP) source systems such as an Enterprise Resource Planning (ERP) system. The cleansing of the data that will give meaningful insight to the business is done with an Extraction Transformation Loading (ETL) tool. Data warehouse stores information from operational systems in online analytical processing (OLAP) cubes. Data warehouse OLAP engine provides means for users to access and analyse the data stored in the data warehouse.

2.4 Data Mining

Data mining application has just been made popular due to the recent classified NSA document made public by a former NSA contractor, Edward Snowden. The allegation made was that the US government has been secretly using data mining to spy and eavesdrop on their citizens and foreign governments through a clandestine program called PRISM where NSA taps directly to the servers of some of the biggest companies in the world such as Google, Microsoft, Yahoo, Apple and collect and store meta data of emails, phone calls, chat conversations, documents. The aim of the program is for the NSA to use predictive analytics to counter terrorism.

Data Mining (DM) or Knowledge Discovery of Data (KDD) refers to extracting or mining knowledge from large amounts of data which is shown in figure 3 below (Kamber et al., 2006: 5). Another meaning of data mining includes Knowledge mining from data, knowledge extraction, data/pattern analysis, data archaeology and data dredging.

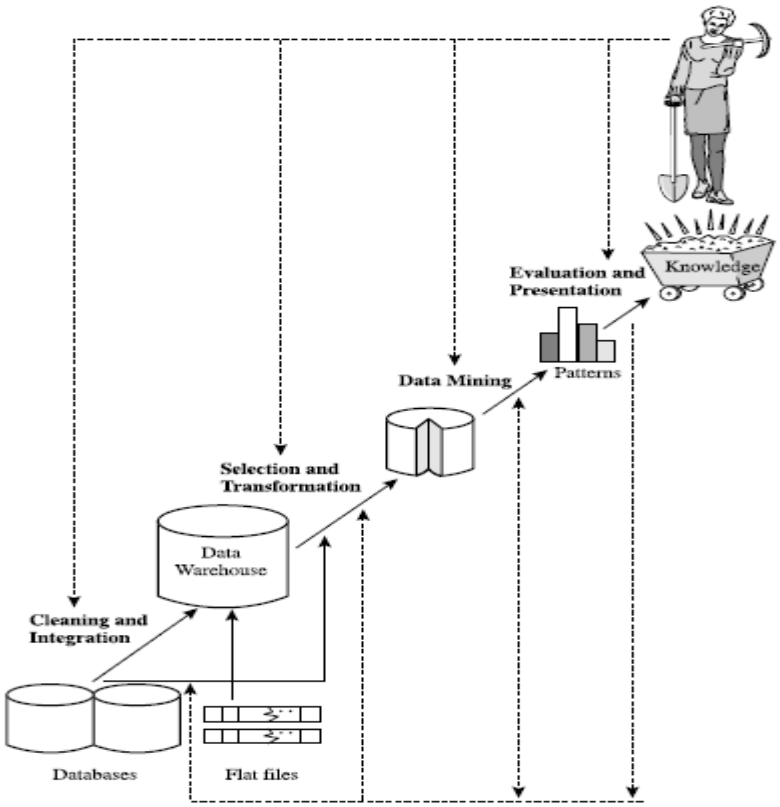

Figure 6 - Data mining as a step in the process of Knowledge discovery (Kamber et al., 2006: 6)

No	Steps	Descriptions
1	Data Cleaning	to remove noise and inconsistent data and redundant data
2	Data integration	where multiple data sources may be combined
3	Data selection	where data relevant to the analysis task are retrieved from the database
4	Data transformation	where data are transformed or consolidated into forms appropriate for mining by performing summary or aggregation operations
5	Data mining	an essential process where intelligent methods are applied in order to extract data patterns
6	Pattern evaluation	to identify the truly interesting patterns representing knowledge based on some interesting-

		ness measures
7	Knowledge presentation	where visualization and knowledge representation techniques are used to present the mined knowledge to the user

Table 1 - Steps in the data mining process (Kamber et al., 2006: 7)

The whole process of presenting the knowledge to the end user involves data preprocess which includes data cleaning, data integration, data selection and data transformation as described in steps 1 – 4. Steps 5 – 7 uncover insights and meaningful meaning of the data and present it to the end user through visualization.

A typical DM system involves databases, data warehouses and information management tools or reporting and analytics tools as shown in Figure 4. Data stored in databases, data warehouses and other external repositories of information are cleaned, transformed and transferred into the data mining engine which then applied series of mathematical algorithms and present the data as required by the end user via front end reporting tools for visualization and analysis.

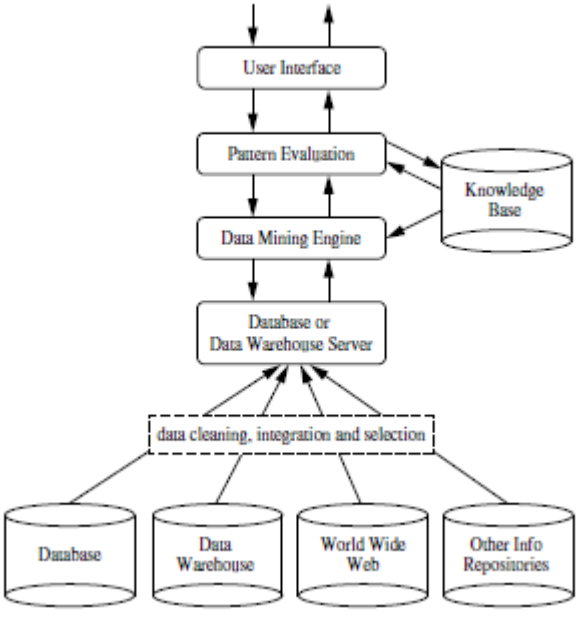

Figure 7 - Architecture of a typical data mining system (Kamber et al., 2006: 8)

The concept of data mining is about finding patterns within business and scientific data and noted for handling large volumes of data to assist in the automation of the knowledge discovery process (Khan et al., 2008: 582). Disciplines such as statistics, Signal or image processing, artificial intelligence (AI), machine learning, database query tools, econometrics, management science, domain-knowledge-based numerical and analytical methodologies, nonlinear dynamic and stochastic systems have all contributed to Data Mining (Khan et al., 2008: 582).

Data mining tasks can be modelled to descriptive and predictive in nature as shown in figure 6. Descriptive data mining includes clustering, summarization, association rules and sequence analysis while predictive involves classification, regression, prediction and time series analysis (Dunham M.H, 2003: 5).Predictive data mining topic will be the primary focus of this thesis report.

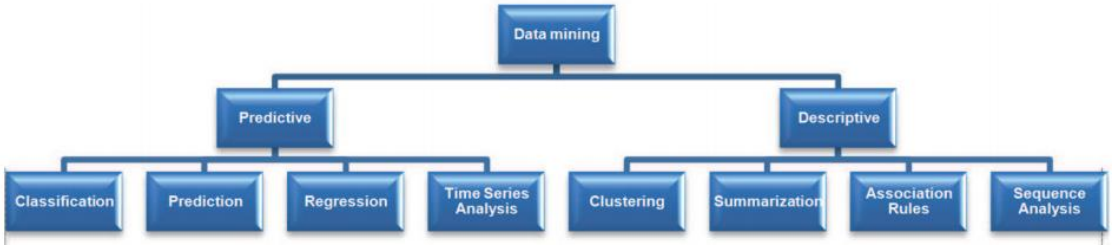

Figure 8 - Data mining models and tasks (Dunham M.H, 2003: 5)

2.5 Business Intelligence

Business intelligence (BI) is a broad category of applications and technologies for gathering, storing, analysing and providing access to data to help enterprise users make better business decisions. Business intelligence applications include the activities of decisions support systems, query and reporting, online analytical processing (OLAP), statistical analysis, forecasting, and data mining (Luca Rossetti, 2006).

Business intelligence applications can be of the following types: a) Mission- critical and integral to enterprise operations or occasional to meet a special requirement; b) enterprise-wide or local to one division, department or project; and c) Centrally initiated or driven by user demand.

Figure 7 shows how data is extracted from source systems to front end BI applications

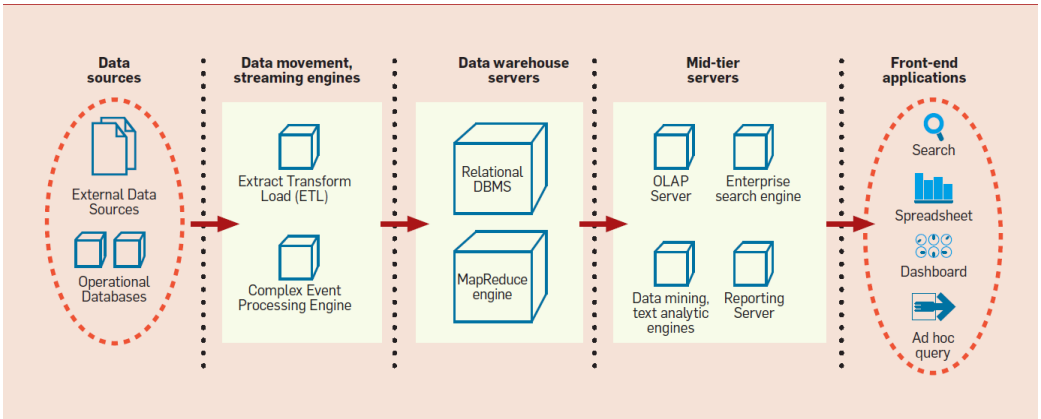

Figure 9 - An overview of Business Intelligence Technology (Surajit Chaudhuri et al., 2010:93)

As Figure 7 demonstrates, data which comes from multiple source systems are transferred and stored in OLAP cubes in data warehouse by an ETL tool and made available to the users through front end reporting applications. ETL tool provides mechanism for data extraction from source systems, data transformation by enriching it with new business logic and cleansing of the data, and finally loading the data into data warehouse OLAP cubes.

The process described above shows how business intelligence applications are created and made available to users for decision making purposes through visualized front end applications.

2.6 In-Memory computing

Predictive analytics in real-time is now feasible due to the ability of in-memory computing to find patterns and process complex data streams in real time. The combination of real time data streams and predictive analytics also known as processing that never stops has the potential to deliver significant competitive advantage for business (Intel IT Center, 2013).

In-Memory computing BI enables data to be stored in random access memory (RAM) rather than on a physical disk as is the case of traditional DW. This storage of data in RAM enables faster access of data, boosts query performance and allows business intelligence and analytical applications to support faster decision-making as a result of near real time capabilities.

Presently, In-memory computing is becoming popular due to the following reasons summarized by (Chaudhury et al, 2011:93). “First, the ratio of time to access data on disk vs. Data in memory is increasing. Second, with 64-bit operating systems becoming common, very large addressable memory sizes (for example, 1TB) are possible. Third, the cost of memory has dropped significantly, which makes servers with large amounts of main memory affordable.”

Also, (Yellowfin, 2011:3) stated that in-memory computing benefits include: first, dramatic performance improvements. In Practice, it means that Users are querying and interacting with data in-memory which is significantly faster than accessing data from disk. The second advantage of in-memory includes cost effective alternative to data warehouses. This is especially beneficial for midsize companies that may lack the expertise and resources to build a data warehouse. The in-memory approach provides the ability to analyze bigger data sets, and it is much simpler to set up and administer. Consequently, IT is not burdened with tasks that take time to perform performance tuning which is typically required by data warehouses.

The third advantage of in-memory computing use is the ability to discover new insights. It means that business users now have self-service access to the right information, coupled with rapid query execution, which allows the delivery of new levels of insight required to optimize business performance, without getting the IT ‘bottleneck’ (ibid).

Based on these characteristics benefits, a typical scenario for the application of in-memory computing is the retail sector whereby management can instantly learn what product a customer has purchased using real-time business insight by analysing the operations as are unfolding, rather than after waiting for a few days or weeks as is the case of using traditional data warehouse for business intelligence. In case of in-memory computing, it is possible because the information can be transferred from source system to data warehouse or front end reporting applications instantly as it is created.

In today’s fast pace business environment, In-memory computing is capable of resolving some of the challenges that users of data warehouse reporting are currently experiencing as depicted in Figure 8 below where all the bottle necks in traditional data warehouse have been removed.

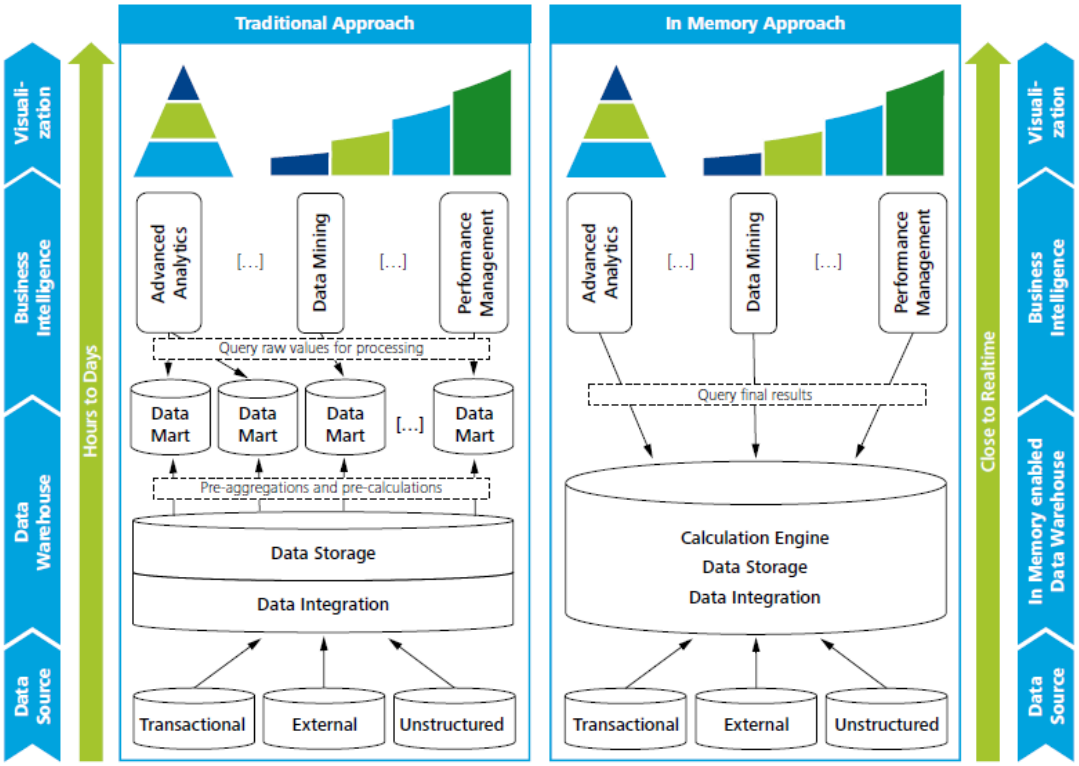

Figure 10 - Traditional data warehouse vs In-memory data Computing technology (In memory computing, the holy grail of analytics, 2013:4)

For this research project, SAP HANA in-memory database technology will be used for the model development. Figure below depicts SAP HANA PAL with its native algo-rithms.

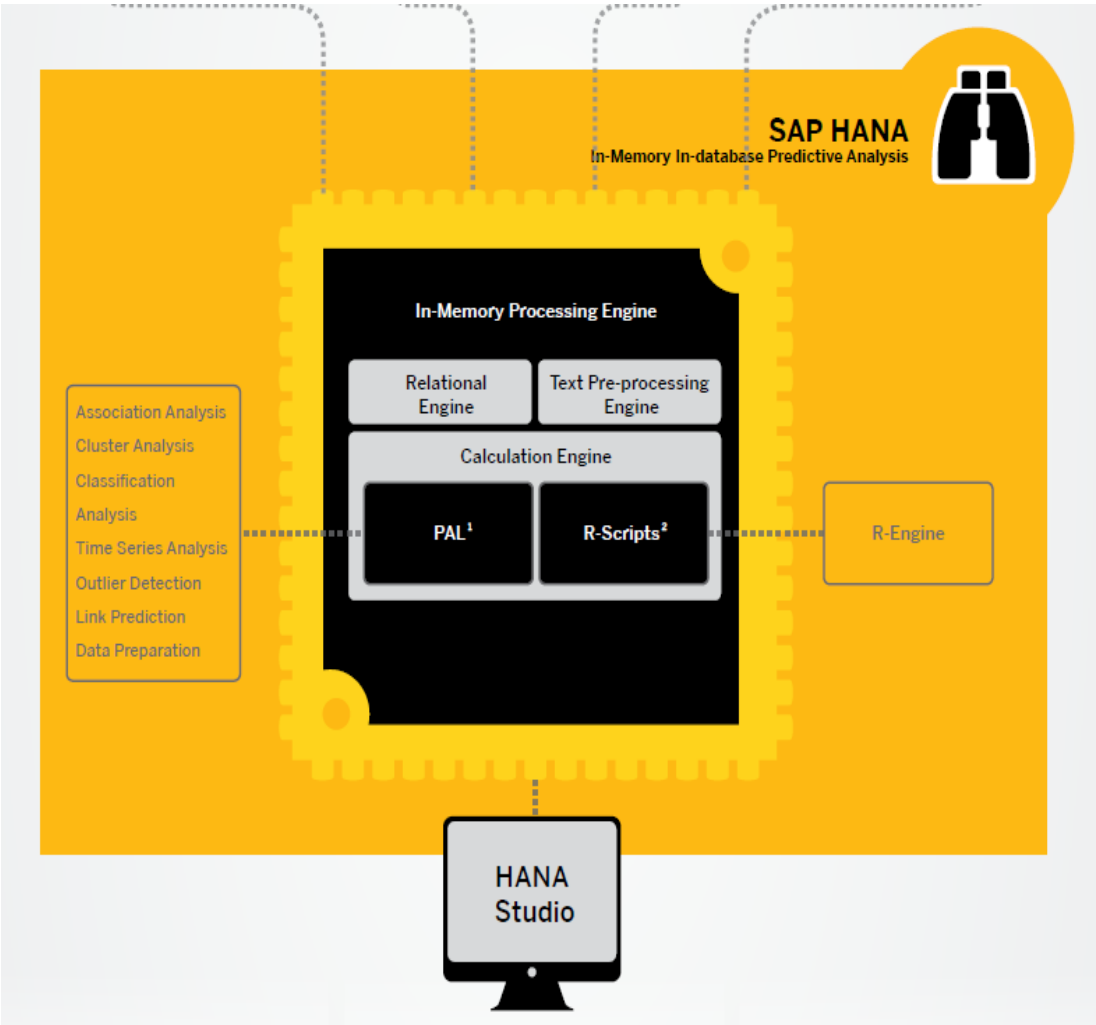

Figure 11 - SAP HANA In-memory database and PAL (SAP HANA Predictive Analysis Library, 2013)

2.7 Predictive Analytics

Predictive analytics (PA) is a branch of business intelligence that applies disciplines such as probability and statistics, machine learning, artificial intelligence, and computer science to business problems (Collete L, The Power of predictive analytics, 2006:12). It uses historical or present data to predict future occurrence of events and unearth hidden meaning of data. In retail banking (Lamonth J, Ph.D, Predictive analytics: an asset to retail banking worldwide, 2005 VOL 14, Issue 10), it allows organizations to access risk and opportunities using historical data to construct models characteristics of a group of customers with their financial behaviour.

Collete (The power of predictive analytics, 2006:1) stated that predictive models analyse past performance to predict future behaviour and this is done by analysing historical and transactional data to isolate patterns and predict an outcome. A typical example is customer churning (ibid).

Predictive analytics is divided into the following, predictive model, descriptive model and decision model. Descriptive model identifies many different relationships in data and classify them into groups while predictive models analyses past performance to predict future behaviour (ibid).
Finally, decision models which is the most advance level of predictive analytics predicts the outcomes of complex decisions in much same way that predictive models predict customer behaviour.

The business cases of predictive analytics can be seen in figure 10 below.

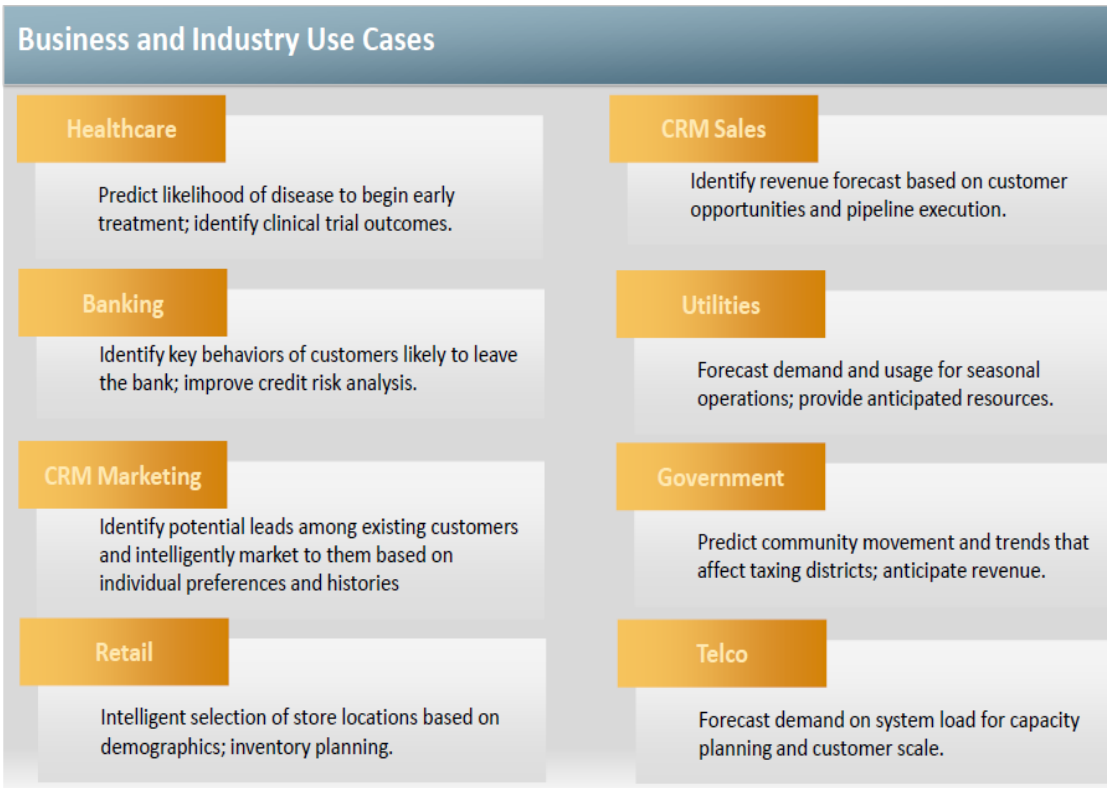

Figure 12 - Predictive analytics business and industry use cases

3 Predictive analytics algorithms

This chapter introduces the different predictive analytics algorithms and their relationship in making predictions.

3.1 Regression analysis

Regression analysis algorithm is commonly used and applied in making forecast, estimation and predictions about future occurrences. It is used to ascertain the relationship between one or more variables and the causal effect they have on one another in order to gain information about one of them through knowing the values of the other, for example understanding the effect of a price increase upon demand or the effect of changes in the money supply upon the inflation rate (Alan Sykes, An introduction of regression analysis, 1996:1). Other example of regression analysis includes the effect of war on high prices of oil (Stephen L., War and its effect on oil prices (2011) and the increase of air ticket prices on popular holiday destinations during summer period (Rebecca B., Cost of summer getaways hit as air ticket rise (2012).

In regression analysis, variables can be independent which is sometimes called predictor or dependent which can be called response variable. Predictor variables are those that can either be set to a desired value (controlled) or else take values that can be observed without any error (Turkman K.F., Linear Regression, 2012:13). Predictor variables can be defined as those variables that can either be set to a desired value (controlled) or else take values that can be observed without any error (ibid).

A regression analysis with a single variable is called simple regression and if two or more variables occur; the regression analysis is called multiple regression analysis (Kerby Shedden, 2013:1).

3.1.1 Simple linear regression analysis

Simple linear regression with a single predictor variable is called simple linear regression and can be defined as (Jiawan H and Micheline K., 2006:355):

$y = b + wx$

Where the variance of y is assumed to be constant, and b and w are regression coefficients specifying the Y – intercepts and slope of the line (Jiawan H and Micheline K., 2006:355):

The table below shows the relationship between Salary earned and number of years of work experience. When plotted in a scatter plot diagram, there is a direct relationship between years of experience in X axis and salary earned annually in Y axis. As the number or years of work experience increases so is the amount of salary earned.
In reality, this is not always the case as there are lots of variables that affect the amount of salary earned.

Number of work experience years	Annual Income
3	30000
8	57000
9	64000
13	72000
6	36000
11	43000
21	90000
1	20000
16	83000

Figure 13 - Number of work experience versus Annual income

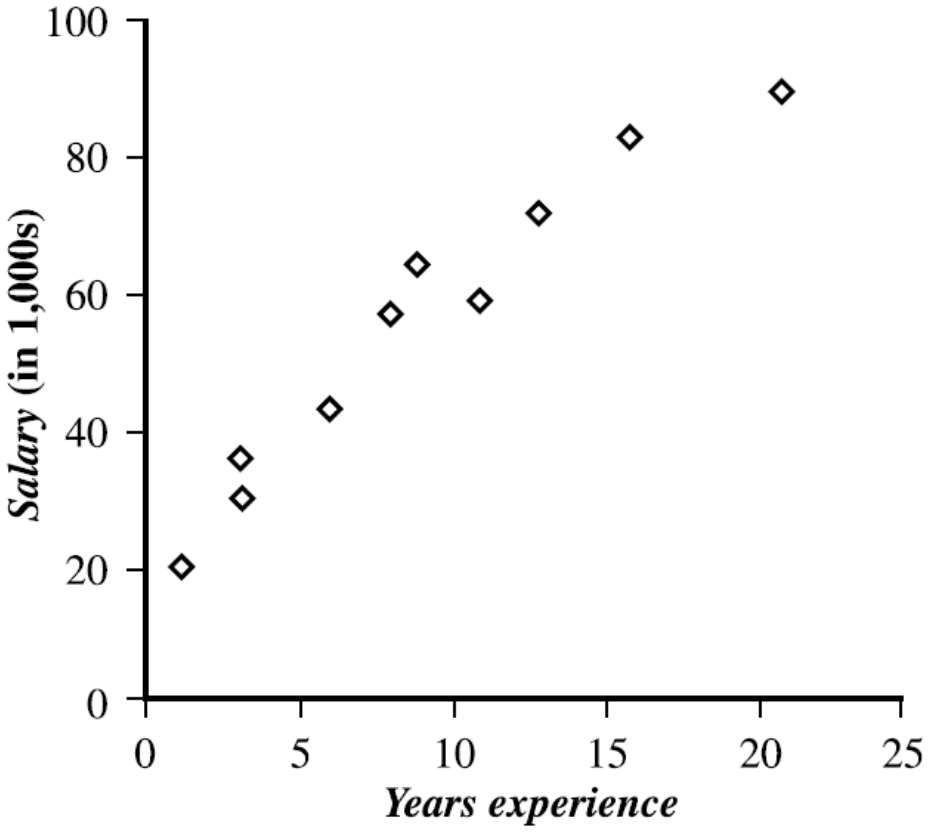

Figure 14 - simple linear regression

3.1.2 Multiple linear regression analysis

Multiple regression analysis is an extension of simple linear regression as it involves more than one predictor variable (Jiawan H and Micheline K., 2006:357). For example figure 9 below includes additional variable called number of years of education.

An example of a multiple linear regression model based on two predictor attributes or variables, *A*1 and *A*2, is

$y = w0+w1x1+w2x2$

Where $x1$ and $x2$ are the values of attributes *A*1 and *A*2, respectively, in X.

The method of least squares shown above can be extended to solve for $w0$, $w1$, and $w2$. The equations, however, become long and are tedious to solve by hand. Multiple regression problems are instead commonly solved with the use of statistical software packages, such as SAS, SPSS, and S-Plus

Number of work experience years	Annual Income	Number of years of Education
3	30000	5
8	57000	4
9	64000	4
13	72000	5
6	36000	4
11	43000	4
21	90000	7
1	20000	3
16	83000	5

Figure 15 - work experience versus Annual income vs education

3.2 Neural Networks

Neural Networks also known as Artificial neural network (ANN) is a predictive analytics algorithm technique modelled after cognitive system and neurological functions of the brain and is capable of predicting new observations on the same or other variables by using artificial intelligence to learn from new and existing data (Statsoftsa, 2013).

In other words, neural network is a simplified model of the biological nervous system and work and draws its motivation from the kind of computing performed by the human brain (Reshma et al., 2013:1). Neural network have been applied successfully to problems in the field of pattern recognition, image processing, data compression, forecasting and optimization (ibid).

Artificial neural networks is made of an artificial neurons modelled after natural biological neurons as seen in the figure 14 below. As stated by (Gershenson C., 2003:1), Natural neurons receive signals through synapses located on the dendrites or membrane of the neuron and these neurons are activated and emits signals through the axon whenever the signals received are strong enough or surpasses certain threshold.

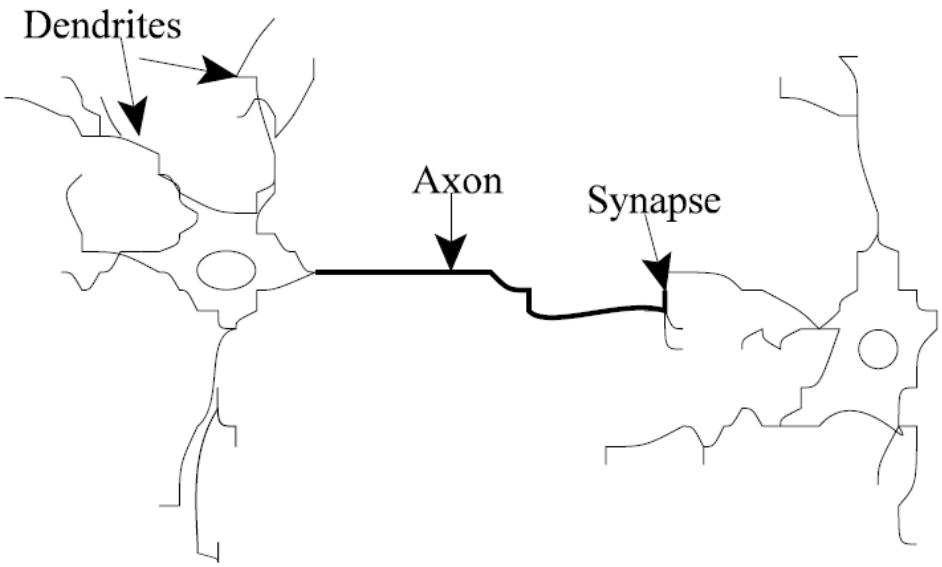

Figure 16 - Natural neurons

ANN consists of inputs or synapses which are multiplied by weights or strengths of the respective signals. The activation of the neuron is determined by a mathematical function as shown in figure 15 (ibid). Patterns are presented to the network via the input layer, which communicates to one or more hidden layers where the actual processing is done via a system of weighted connections.

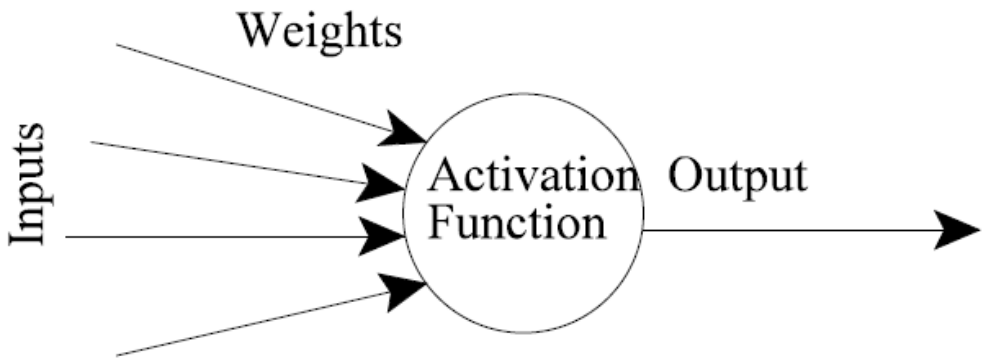

Figure 17 - Artificial neuron

ANNs are based on the combination of neurons, connections and transfer functions with various learning algorithms any layout methods for the neurons and their connections. The limitation of ANN is that it acts essentially as a black box that performs assigned tasks without the user having control of the output. The information stored in

neural networks is a set of numerical weights and connections that provides no direct clues as to how the task is performed or what the relationship is between the inputs and outputs. This limits the usage and acceptance of ANN since in many applications science and engineering it is demanded to use techniques based on analytical functions that can be understood and validated (Tzeng et al., 2005:1).

Examples of neural networks applications can be seen in the table below.

Application	**Description**
Detection of medical phenomena	A variety of health-related indices (e.g., a combination of heart rate, levels of various substances in the blood, respiration rate) can be monitored. The onset of a particular medical condition could be associated with a very complex (e.g., nonlinear and interactive) combination of changes on a subset of the variables being monitored. Neural networks have been used to recognize this predictive pattern so that the appropriate treatment can be prescribed.
Stock market prediction	Fluctuations of stock prices and stock indices are another example of a complex, multidimensional, but in some circumstances at least partially-deterministic phenomenon. Neural networks are being used by many technical analysts to make predictions about stock prices based upon a large number of factors such as past performance of other stocks and various economic indicators.
Credit assignment	A variety of pieces of information are usually known about an applicant for a loan. For instance, the applicant's age, education, occupation, and many other facts may be available. After training a neural network on historical data, neural network analysis can identify the most relevant characteristics and use those to classify applicants as good or bad credit risks.

Monitoring the condition of machinery	Neural networks can be instrumental in cutting costs by bringing additional expertise to scheduling the preventive maintenance of machines. A neural network can be trained to distinguish between the sounds a machine makes when it is running normally ("false alarms") versus when it is on the verge of a problem. After this training period, the expertise of the network can be used to warn a technician of an upcoming breakdown, before it occurs and causes costly unforeseen "downtime."
Engine management	Neural networks have been used to analyze the input of sensors from an engine. The neural network controls the various parameters within which the engine functions, in order to achieve a particular goal, such as minimizing fuel consumption.

Table 2 - Examples of Neural networks (Statsoft, 2013)

The different neural networks algorithms include:

3.2.1 Feed-forward network

A feed-forward network is a non-recurrent network which contains inputs, outputs, and hidden layers; the signals can only travel in one direction. Input data is passed onto a layer of processing elements where it performs calculations. Each processing element makes its computation based upon a weighted sum of its inputs. The new calculated values then become the new input values that feed the next layer.

This process continues until it has gone through all the layers and determines the output. A threshold transfer function is sometimes used to quantify the output of a neuron in the output layer. Feed-forward networks include Perceptron (linear and non-linear) and Radial Basis Function networks. Feed-forward networks are often used in data mining (Saed Sayad, 2013).

3.2.2 Feed-backward network

A feed-back network has feed-back paths meaning they can have signals traveling in both directions using loops. All possible connections between neurons are allowed. Since loops are present in this type of network, it becomes a non-linear dynamic system which changes continuously until it reaches a state of equilibrium. Feed-back networks are often used in associative memories and optimization problems where the network looks for the best arrangement of interconnected factors (ibid).

3.3 Time Series

In early Babylonian astronomy, time series have been used to predict numerous astronomical events using the relative positioning of stars through observations of the movement of the planets. And this has formed the basis of planetary laws put forward by Johannes Kepler (G Kirchgassner and J. Wolters, Introduction to Modern Time Series, 2007:2).

The analysis of time series helps to detect regularities in the observations of a variable and derive ‘laws’ from them, and/or exploit all information included in this variable to better predict future developments. The basic methodological idea behind these procedures, which were also valid for the Babylonians, is that it is possible to decompose time series into a finite Number of independent but not directly observable components that develop regularly and can thus be calculated in advance. For this procedure, it is necessary that there are different independent factors which have an impact on the variable (ibid).

G. Kirchgassner and J. Wolters (2007:2) defined time series analysis as a set of quantitative observations of data collected over time arranged in chronological order usually daily values, weekly values, monthly values, quarterly values and yearly values. Usually the observations are made repeatedly over 50 or more time periods. The observations can be from a single case or an aggregated score from many cases tracked over a considerable time for example, the scores might represent the daily number of temper tantrums of a two year old, the weekly output of a manufacturing plant, the monthly number of traffic tickets issued in a municipality, or the yearly GNP for a developing country (Hery Mulyana, 2011:1).

The intention is to determine if there is pattern in the data collected to date with the aim to make predictions of future developments or forecast.

The technique that is used in time series data to make forecast or prediction is the exponential smoothing which assigns exponentially decreasing weights over time or as the observations get older unlike moving average method which compute the average of the most recent data values for the series and using the average to forecast the value of the time series for the next period. It can be said that exponential smoothing method gives more weight to recent observations in forecasting than the older observations.

The three types of exponential smoothing methods are the single exponential smoothing (SES) double (Holt's) exponential smoothing (DES) and triple (winter's) exponential smoothing.

3.3.1 Single exponential smoothing

This is also known as simple exponential smoothing and it is a suitable to model time series when there is no trend or seasonality in the data. When the data exhibits either an increasing or decreasing trend over time, simple exponential smoothing forecasts tend to lag behind observations.
The formula for single exponential smoothing is as follows (Marzena et al., 2013:31):

$F_{t+1} = F_t + \alpha(y_t - F_t)$ or
$F_{t+1} = \alpha y_t + (1 - \alpha) F_t$
Where:
Ft+1= forecast value for period t + 1
yt = actual value for period t
Ft = forecast value for period t
α = alpha (smoothing constant)

When applied recursively to each successive observation in the series, each new smoothed value (forecast) is computed as the weighted average of the current observation and the previous smoothed observation; the previous smoothed observation was computed in turn from the previous observed value and the smoothed value before the previous observation, and so on.

3.3.2 Double exponential smoothing

Double exponential smoothing also called smoothing with trend is used for forecasting when there is trend in the data but no seasonality. This smoothing works much like the single exponential smoothing except that the components is updated with trend.

The formulas for double exponential smoothing are (Marzena et al., 2013:37):

$C_t = \alpha y_t + (1 - \alpha)(C_{t-1} + T_{t-1})$

$T_t = \beta(C_t - C_{t-1}) + (1 - \beta) T_{t-1}$

$F_{t-1} = C_t + T_t$

Where:

yt = actual value in time t

α = constant -- process smoothing constant

β = trend -- smoothing constant

Ct = smoothed constant -- process value for period t

Tt = smoothed trend value for period t

F_{t+1} = forecast value for period t + 1

t = current time period

3.3.3 Triple exponential smoothing

Triple exponential smoothing also known as the Winters method is a refinement of the popular double exponential smoothing model but adds another component which takes into account any seasonality or periodicity in the data.

As with simple exponential smoothing, in triple exponential smoothing models past observations are given exponentially smaller weights as the observations get older. In other words, recent observations are given relatively more weight in forecasting than the older observations. This is true for all terms involved - namely, the base level L_t, the trend T_t as well as the seasonality index s_t.

There are four equations associated with Triple Exponential Smoothing.

$L_t = a.(x_t/s_{t-c})+(1-a).(L_{t-1}+T_{t-1})$

$T_t = b.(L_t-L_{t-1})+(1-b).T_{t-1}$

$s_t = g.(x_t/L_t)+(1-g).s_{t-c}$

$f_{t,k} = (L_t+k.T_t).s_{t+k-c}$

where:

L_t is the estimate of the base value at time t. That is, the estimate for time t after eliminating the effects of seasonality and trend.

a - representing alpha - is the first smoothing constant, used to smooth L_t.

x_t is the observed value at time t.

s_t is the seasonal index at time t.

C is the number of periods in the seasonal pattern. For example, c=4 for quarterly data, or c=12 for monthly data.

T_t is the estimated trend at time t.

b - representing beta is the second smoothing constant, used to smooth the trend estimates.

g - representing gamma is the third smoothing constant, used to smooth the seasonality estimates.

$f_{t,k}$ is the forecast at time the end of period t for the period t+k.

3.4 Decision Trees

Decision trees are a simple, but powerful form of multiple variable analyses and provide unique capabilities to supplement complement and substitute for the following (Berry de Ville, 2006:1):

- Traditional statistical forms of analysis such as multiple linear regression
- A variety of data mining tools and techniques such as neural networks.
- Recently develop multidimensional forms of reporting and analysis found in the field of business intelligence.

Decision tree is a popular predictive analytics tool because of the way data is presented visually. The visual presentation of data by decision trees makes it easy to read, understand and assimilate information.

Berry and Linoff (2006:166) defined decision tree as a tree structure that can be used to divide up large collection of records into successively smaller sets of records by applying a sequence of simple decision rules. The members of each successive division resulting sets become more and more similar to one another. This is akin to the familiar division of living things into kingdoms, phyla, classes, orders, families, genera and species (ibid).

Decision trees are powerful and popular data mining tool for classification and predictions. It is a tree in which each branch node represents choice between a number of alternatives, and each leaf node represents a classification or decision (ibid).

The model for decision tree consists of a set of rules for dividing a large datasets into smaller, more homogeneous datasets with respect to a particular target variable (ibid)

Decision tree consist of nodes that form a rooted tree. This is a directed tree with a node called the root which has no incoming edges. Other nodes have exactly one in-

coming edge. A node that has outgoing edges is called an internal or test node. All other nodes are called leaves which can also be called terminal or decision nodes.
There are numerous applications of decision trees. An example as shown below is the owner of an ice cream stand that wants to know what makes people will by ice cream (Padraic G. Neville, 1999:2).

The observed data in the example showed that forty percent of people buy ice cream. This is represented in the root node of the tree at the top of the diagram. The first rule splits the data according to the weather. Unless it is sunny and hot, only five percent buy ice cream. This is represented in the leaf on the left branch. On sunny and hot days, sixty percent buy ice cream. The tree represents this population as an internal node that is further split into two branches, one of which is split again.

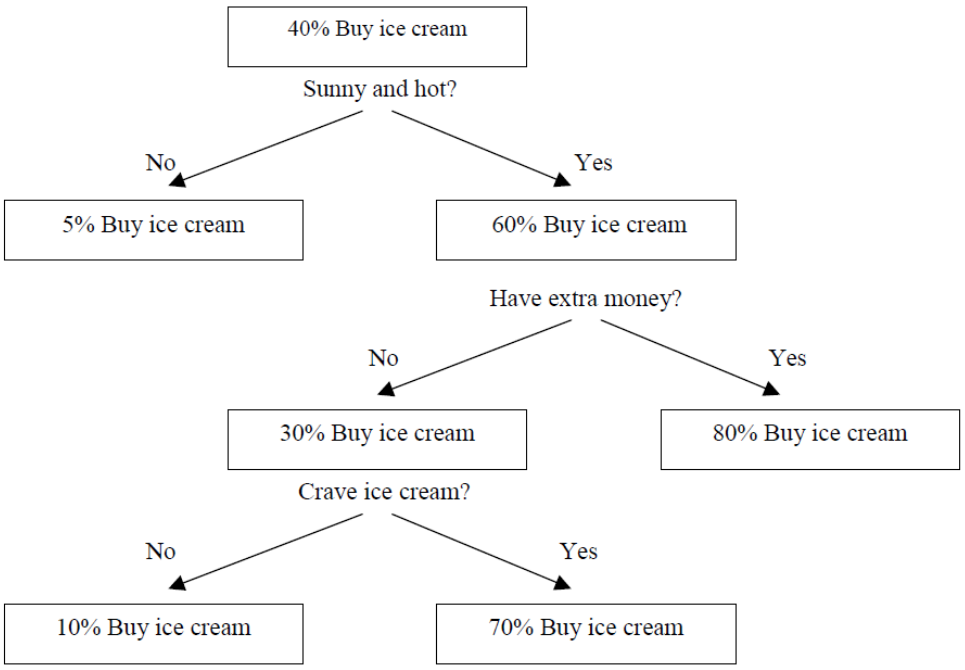

Figure 18 - Decision tree (Padraic G. Neville, 1999:2)

According to the tree, people almost never buy ice cream unless the weather cooperates and either (1) they have some extra money to spend or (2) they crave the ice cream and presumably spend money irresponsibly or figure out another way of buying it (ibid).

3.5 Clustering

Clustering is a descriptive data mining algorithm that categorizes observations in a database into groups called clusters. This algorithm categorizes the total number of cases into a smaller number of groups such that the cases within each group are similar to each other, but dissimilar to the cases in other groups (Shiraj Khan et al., 2008:584).

In clustering analysis, objects or observations in a specific cluster share many characteristics, but are very different from objects or observations belonging to other cluster.

Clustering analysis is widely use in customer market analysis, where customers are segmented based on different factors or criteria variables for example their income or level of education in order to target them for better pricing strategies. The segmentation of customers is a standard application of cluster analysis, but it can also be used in different, sometimes rather exotic, contexts such as evaluating typical supermarket shopping paths or deriving employers branding strategies (Larson et al. 2005:1).

Cluster models resulting from cluster analysis have traditionally been used quite extensively in marketing applications to help characterize groups of similar consumers. The ability to better understand these groups can lead to more effective messaging and new product development efforts.

Some of the most popular methods used in clustering analysis include K-Means algorithm and Kohonen self-organizing map (SOM).

3.5.1 K-Means

K-means clustering is a method that attempts to assign a set of *n* observations into a *k* number of clusters where each observation is allocated to the cluster with the nearest center point. Therefore each observation can only belong to one cluster. The center point of a cluster is the mean value for all the observations (Correa et al, 2012:2).

Lloyd's algorithm commonly known as standard algorithm or K-Means algorithm is one of the heuristic algorithms that are used to reach an optimum assignation of the observations to the cluster (ibid). This method is usually iterative technique to reach optimum clustering.

Correa et al (2012:2) described that the whole process can be divided in two steps. Assignment is the first step, where each observation is assigned to the cluster with the

closes center point. The new center points are calculated based on the observations that formed the cluster at the end of step one in the second step. This process is repeated until the clusters remain unchanged. The goal of this process is to find the best fit to the data, minimizing the within –cluster sum of squares (ibid).

3.5.2 Kohonen clustering Method

Kohonen is an unsupervised and competitive commonly used clustering method. It comes from a self-organizing map (SOM) that is well-known dimension reduction method. Kohonen clustering method have some similarities with K-Means procedure such as the way that new observations are assigned to the clusters and that both methodologies are heuristic process, but the process as a whole is very different (Correa et al.,2012:3).

In self-organizing cluster algorithm, clusters are determined when the nearest cluster called the wining cluster are moved closer towards the training case or observation. The amount of the movements depends of the distance between the winning cluster and the training case, and they decreased throughout the process by means of the learning rate (ibid).

A neural network seen below is the basis of SOM cluster. The input layer is compounded by the k variables (characteristics) of each of the N observations. There is no connection between the output nodes but every node has a connection with all the input nodes.

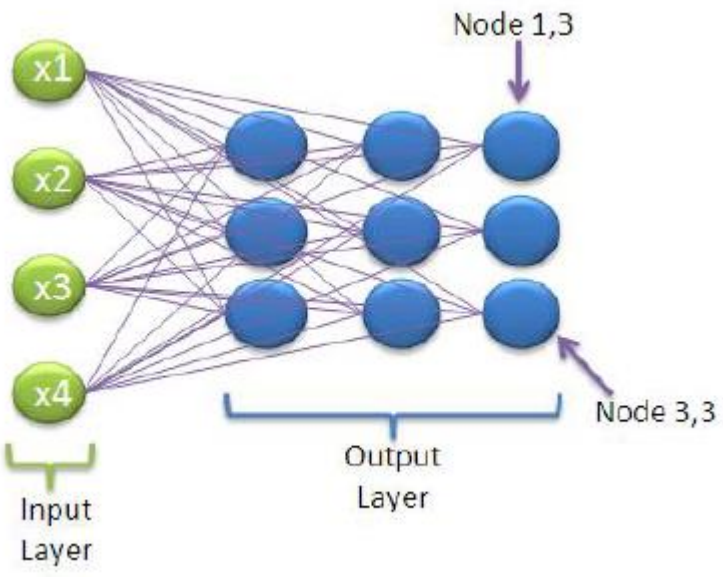

Figure 19 - Kohonen Neural Network ((Correa et al, 2012:3)

3.6 Association Rules Analysis

Association rules analysis or market basket analysis is a popular tool for mining very large scale commercial database. The rules attempts to describe regions of relatively high density in a very large commercial database.

Market basket analysis concerns the analysis of various subsets of items taken from a population of items. The subsets of concern, or rules, are those identified as having a minimum value of confidence, support or lift. An example rule is denoted as *A->B*, where A is referred to as the rule's antecedent condition and B the consequent. The rule is interpreted as "If A occurs in the market basket, then B also occurs in the market basket." (Sanford Gayle, 2000:1).

For example, consider the sales database of an on-line retailer e.g. Amazon, where the objects represent customers and the attributes represent items or products. The rules to be discovered are the set of items or products most frequently bought together by the customers. An example could be that

- 15% of the people who buy Dorian Pyle's Data Preparation for Data Mining also buy Data Mining Techniques by Berry and Linoff."
- 60% of the customers who buy milk also buy bread and eggs
- 80% of the time that a specific brand of toaster is sold, customers also buys a set of kitchen gloves and matching cover sets.

Also, movie rental such as Netflix uses market basket analysis to recommend movies for their customers for example customers that watches "Lord of the rings also watches the hobbit".

Some of the applications of market basket analysis include:

- Analysis of purchases made with a credit card.
- Analysis of telephone calling patterns
- Analysis of medical history
- Analysis of telecom service purchase
- Identification of fraudulent claims

3.7 Classification

Classification is a data mining technique that is used for discovering classes of unknown data. Classification follows supervised learning technique whereby a set of examples whose label are known are given to the algorithm.

Classification consists of predicting a certain outcome based on a given input. In order to predict the outcome, the algorithm processes a training set containing a set of attributes and the respective outcome, usually called goal or prediction attribute (Fabricio Voznika and Leonardo Viana, 2007:1).Classification algorithm tries to discover relationships between the attributes that would make it possible to predict the outcome. Next the algorithm is given a data set not seen before, called prediction set, which contains the same set of attributes, except for the prediction attribute – not yet known.

The data classification process: Learning: Training data are analyzed by a classification algorithm. Here, the class label attribute is loan decision, and the learned model or classifier is represented in the form of classification rules. (b) Classification: Test data are used to estimate the accuracy of the classification rules. If the accuracy is considered acceptable, the rules can be applied to the classification of new data tuples. (Kamber et al., 2006: 287)

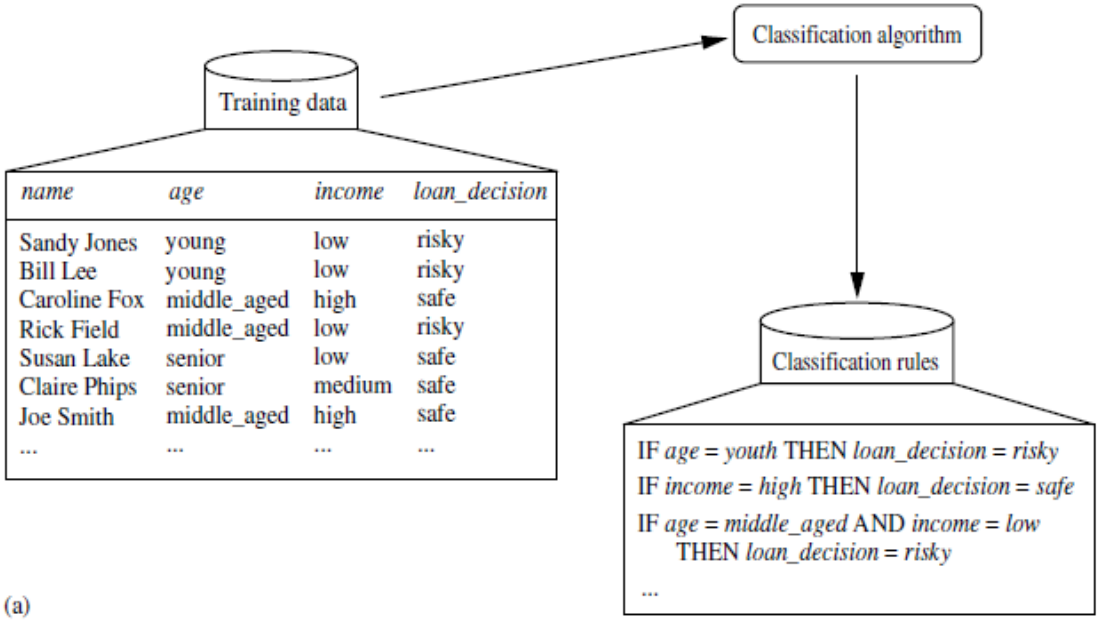

Figure 20 - Classification training data (Kamber et al., 2006: 287)

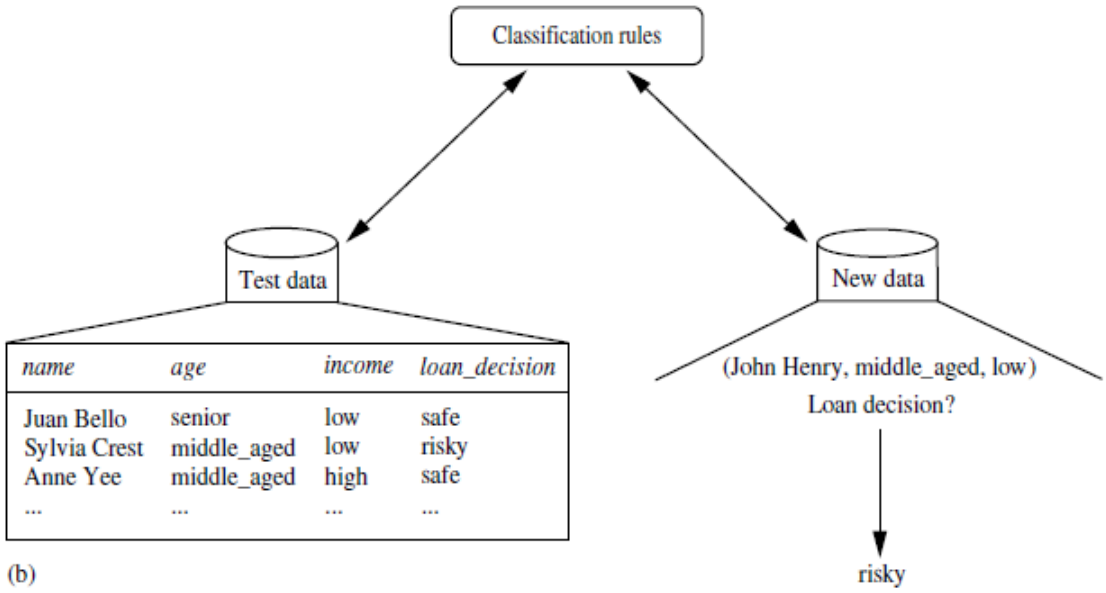

Figure 21 - Classification test data (Kamber et al., 2006: 287)

Classification technique includes rule based classifier, genetic algorithm, Bayesian classifier, k-nearest neighbor, rough sets and fuzzy logic.

3.7.1 Rule Based Classifier

Rule based classifier deals with the discovery of high-level, easy to interpret classification rules of the form IF-THEN (Beniwal and Arora (2012:3). The rules are composed of two parts mainly rule antecedent and rule consequent. The IF part which specifies a set of conditions referring to predictor attribute values is the rule antecedent while the part that satisfies the conditions in the rule antecedent is the THEN.

3.7.2 Bayesian Classification

Bayes Theorem which can also be called Naïve Bayesian classification was first proposed by Thomas Bayes. This classification represents a supervised learning method as well as a statistical method of classification. This algorithm provides practical learning algorithms and prior knowledge and uses the learned knowledge to predict future events (Mihaesu C.2002:1).

They can predict class membership probabilities, such as the probability that a given sample belongs to a particular class. Naive Bayesian classifiers assume that the effect of an attribute value on a given class is independent of the values of the other attributes. This assumption is called class conditional independence.

One of the practical usages of Bayesian classification is email spam filtering because the algorithm can distinguish illegitimate spam mail from legitimate email (ibid).

3.7.3 Genetic Algorithm

Genetic Algorithms (GAs) are part of evolutionary computing which is a growing area of artificially intelligence (AI). GAs is adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. It was inspired by Charles Darwin survival of the fittest evolutionary theory. Genetic algorithm attempts to exploit historical information to direct search into the region of better performance within the search space.

Genetic algorithms are good at taking large, potentially huge search spaces and navigating them, looking for optimal combinations of things, the solutions that would otherwise take a life time to find (RC Chakraborty, 2010:1).

Some of the applications of classification analysis include:

- Credit approval
- Target marketing
- Medical diagnosis
- Treatment effectiveness analysis

3.7.4 K-nearest neighbour

K-Nearest neighbour (KNN from now on) is also called lazy learning algorithm because it doesn't make assumptions on the underlying data distribution. It is also a lazy algorithm. What this means is that it does not use the training data points to do any generalization. In other words, there is no explicit training phase or it is very minimal. (Saravanan Thirumuruganathan, 2007).

KNN classifiers are based on learning by analogy, that is, by comparing a given test tuple with training tuples that are similar to it. The training tuples are described by n attributes. Each tuple represents a point in an n-dimensional space. In this way, all of the training tuples are stored in an n-dimensional pattern space. When given an unknown tuple, a k-nearest-neighbour classifier searches the pattern space for the k training tuples that are closest to the unknown tuple. These k training tuples are the k "nearest neighbours" of the unknown tuple (Kamber et al., 2006: 348).

4 Model Performance Evaluation

This chapter introduces the various statistical methods to evaluate the performance of a predictive model and measures its accuracy. Robert Kunst (2012:58) stated that there are two ways to determine how good a forecast is. Firstly, the predictive accuracy is measured per se and secondly is by comparing various forecasting models. In this thesis paper, the accuracy of the model will be determined by using scientific measuring tools.

4.1 Goodness-of-fit

One of the ways of evaluating the performance of a predictive analytic model is by looking at the accuracy of the model. This is done by comparing the predictive and the actual outcome of the result. When the predicted result outcome is close to the actual result outcome, the model is considered a good fit or has goodness-of-fit. In general, a good predictive model is when the difference between the observed values and the predicted values are small and unbiased (Jim Frost, 2013).

Goodness-of-fit is determined using the Chi-Square distribution method which measures the observed frequency and expected frequency and the formula is (Cohen, H., 2012:459).

$$x^2 = \sum \frac{(F_o - F_e)^2}{F_e}$$

Where

$\sum$ = is taken over all the categories

F_0 = Observed frequency

F_e = Expected frequency

Cohen, Barry (2012:.459) stated that, if the differences between the observed frequencies and the expected frequencies are small, x^2 will be small and that the greater the difference between the observed frequencies and those expected under the null hypothesis, the larger x^2 will be.

The common measures used for measuring the goodness-of-fit are Root Mean Square Error, Relative Square Error and Coefficient of Determination.

4.2 Root Mean Square Error

The Root Mean Square Error (RMSE) (also called the root mean square deviation, RMSD) is a frequently used measure of the difference between values predicted by a model and the values actually observed from the environment that is being modelled. These individual differences are also called residuals or prediction errors, and the RMSE serves to aggregate them into a single measure of predictive power. The calculation for residuals is:

Residuals = Actual value – Predicted value

The RMSE of a model prediction with respect to the estimated variable X_{model} is defined as the square root of the mean squared error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (X_{obs,i} - X_{model,i})^2}{n}}$$

where X_{obs} is observed values and X_{model} is modelled values at time/place *i*.

RMSE values can be used to distinguish model performance in a calibration period with that of a validation period as well as to compare the individual model performance to that of other predictive models.

RMSE compares models whose errors have the same unit. An RMSE with lower value or value close to zero means the model is good and a higher RMSE value denotes a poor model.

4.3 Relative Square Error

Unlike RMSE, the relative squared error (RSE) can be compared between models whose errors are measured in the different units (Saed Sayad, 2013).The root relative squared error is relative to what it would have been if a simple predictor had been used. More specifically, this simple predictor is just the average of the actual values. Thus, the relative squared error takes the total squared error and normalizes it by dividing by the total squared error of the simple predictor. By taking the square root of the relative squared error one reduces the error to the same dimensions as the quantity being predicted.

Mathematically, the root relative squared error E_i of an individual program *i* is evaluated by the equation:

For a perfect fit, the numerator is equal to 0 and E_i = 0. So, the E_i index ranges from 0 to infinity, with 0 corresponding to the ideal (Saed Sayad, 2013).

$$RSE = \frac{\sum_{i=1}^{n}(p_i - a_i)^2}{\sum_{i=1}^{n}(\bar{a}_i - a_i)^2}$$

Where a is actual target and p is predicted target (ibid).

4.4 Mean Absolut error

The mean absolute error is a statistical method used to determine the accuracy of a model when compared to actual and historical observations (Contextuall, 2012).

To determine the model accuracy, each predicted value by the model is compared to the actual value observed and the absolute errors are averaged to develop an estimate of the model's accuracy.

The mean absolute error (MAE) has the same unit as the original data, and it can only be compared between models whose errors are measured in the same units. It is usually similar in magnitude to RMSE, but slightly smaller.

The formula for is:

$$MAE = \frac{\sum_{i=1}^{n}|p_i - a_i|}{n}$$

Where p is predicted target, a is actual target and n is the number of observations (ibid).

4.5 Coefficient of Determination

Coefficient of determination or R Squared is a statistical measure of how close the data are fitted in the regression line and it can be defined as the percentage of the response variable variation of the model. The formula is :

$$R^2 = \frac{Explained\ variation}{Total\ variabtion}$$

The value of R-Squared is always between 0 – 100%:

0% indicates that the model explains none of the variability of the response data around its mean

100% indicates that the model explains all the variability of the response data around its mean (Jim Frost, 2013). Higher values for R-Squared signify that the predictive model is good and that the model fits the data.

4.6 Prediction Error

When making predictions, assumptions are made that the underlying data follows some underlying mathematical model and during training, the training data is fitted into this assumed model to determine the best model parameter that will give minimal error (Ricky Ho, 2012).

These assumptions often lead to two kinds of prediction errors namely error due to bias and error due to variance. Understanding these two types of error can help us diagnose model results and avoid the mistake of over-or under-fitting and help improve the data fitting process to that will result in a more accurate model. Bias and variance can be defined conceptually and graphically (Scott Fortmann-Roe, 2012).

4.7 Conceptually

Biased errors can be defined conceptually in terms of error due to bias and error due to variance.

4.7.1 Error due to bias

Error due to bias is taken as the difference between the expected (or average) prediction of our model and the correct value which we are trying to predict (ibid).This is when the assumed model is fundamentally wrong for example if the output has a nonlinear relationship with the input and the model is assumed to be linear model.

4.7.2 Error due to Variance

Error due to variance is taken as the variability of a model prediction for a given data point. For example, repeating the entire model building process multiple times. The variance is how much the predictions for a given point vary between different realizations of the model (ibid).

4.8 Graphically

Bias and variance can be visualized graphically using bulls-eye diagram. The center of the target is the model that perfectly predicts the correct values and as we move away from the center of the bulls-eye, the prediction gets worse and worse. The model building process is represented as number of separate hits on the target. Each hit represents an individual realization of our model.

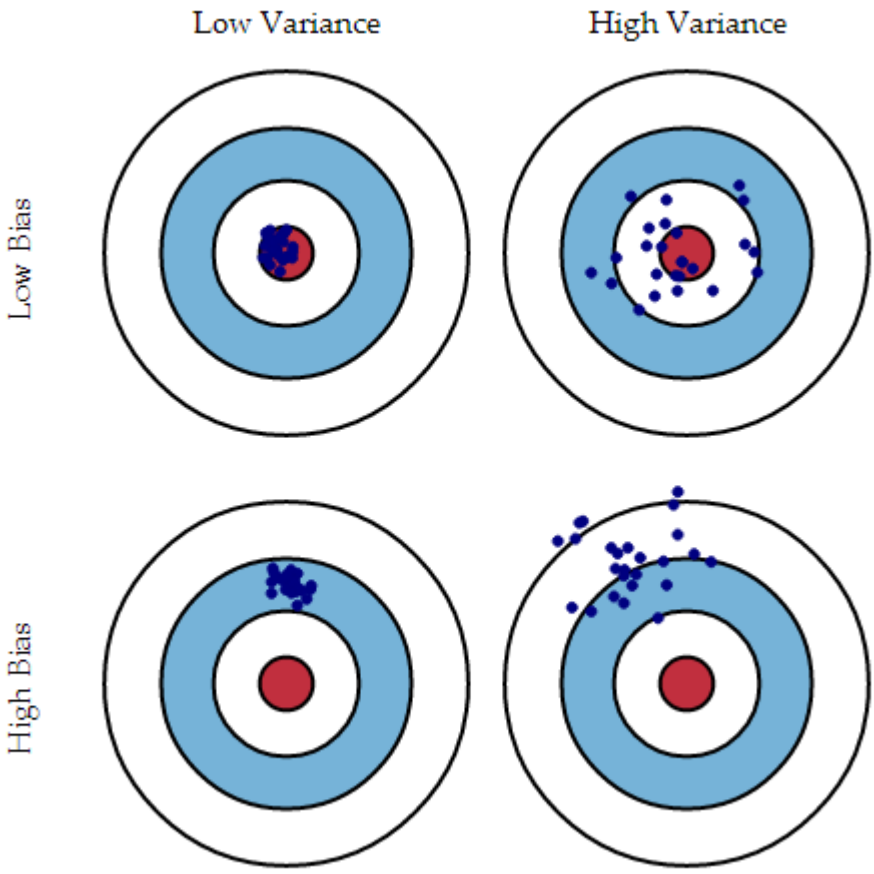

Figure 22 - Graphical illustration of bias and variance (ibid)

5 Predictive model development

This chapter focuses on currency forecasting model development using SAP HANA PAL Time Series algorithm. The model will attempt to compare the forecasted currency exchange rate against the actual exchange rate to determine the accuracy of time series algorithm.

Foreign exchange market popular called Forex, is the biggest market in the world in terms of trading volume with daily trading exceeding more than 4 Trillion USD and the market is dominated by big banks, corporations and private investment funds [38].
There are different factors which make currency exchange rate forecast difficult. Such factors include political, physiological and economic. There are 2 ways of forecasting currency exchange rate. The first one is fundamental analysis which includes purchasing power parity (PPP), relative economic strength and econometric models. The other method is technical analysis which uses past data and mathematical techniques to forecast currency exchange rate. The forecasting model development will base on technical analysis.

5.1 Model development

The overall predictive model development is depicted in figure 18 below. It shows the different phases from data collection to data analysis and results.

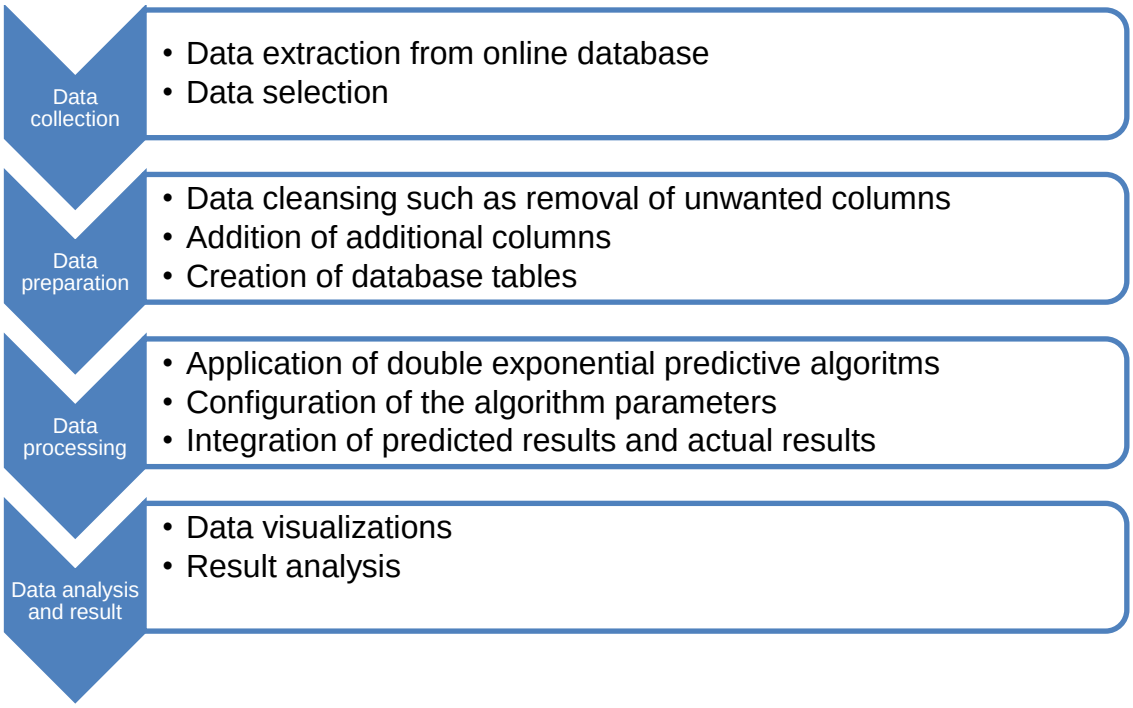

Figure 23 - Model development flow chart

5.2 Data collection

The data for historical exchange rate collected for this research paper was extracted from Norwegian bank online database (Norges-bank 2013). The data is updated at 3:15 PM on a daily bases and has several currency exchange rates. Similar historical exchange rates data can also be collected from other sources. One of the reasons of choosing the exchange rate from Norwegian bank was because it was easy to export to excel without having to do much conversion of the file into the right format.

The data is extracted from the database is more than 13 years of data and it is from the period of 4th January, 1999 to 30th October, 2013. The number of observations of the data which is 3741 records is consistent with Time Series algorithm because at least 52 observations are required to make any meaningful predictions. There is no currency exchange during weekends and public holidays. Figure 23 below represents the data extracted.

	A	B	C	D	E	F	G	H	I	J
1	NB NORGES BANK									
2	8.11.2013	15:03								
3	**Valutakurser, indikative midtkurser**									
4		**Sortert etter valutakode**								
5	Land	Australia	Bulgaria	Canada	Sveits	Tsjekkia	Danmark	Estland	Europeiske Union	Storbritannia
6	Valuta	Dollar	Lev	Dollar	Franc	Koruna	Krone	Kroon	Euro	Pund
7	NOK per:	1 AUD	100 BGN	1 CAD	100 CHF	100 CZK	100 DKK	100 EEK	1 EUR	1 GBP
22	21.10.2013	5,7232	413,74	5,7517	655,12	31,361	108,48		8,0920	9,5594
23	18.10.2013	5,7140	414,08	5,7481	656,12	31,411	108,58		8,0985	9,5823
24	17.10.2013	5,7160	414,92	5,7680	658,31	31,571	108,80		8,1150	9,5741
25	16.10.2013	5,7191	416,02	5,7841	659,31	31,684	109,08		8,1365	9,6187
26	15.10.2013	5,7302	415,07	5,8089	656,74	31,692	108,84		8,1180	9,5918
27	14.10.2013	5,6819	415,81	5,7920	659,36	31,817	109,03		8,1325	9,5868
28	11.10.2013	5,6771	415,61	5,7673	660,10	31,834	108,97		8,1285	9,5742
29	10.10.2013	5,7089	418,75	5,8279	665,26	32,081	109,79		8,1900	9,6569
30	9.10.2013	5,6717	414,66	5,7809	658,65	31,682	108,72		8,1100	9,5818
31	8.10.2013	5,6454	413,23	5,7753	657,93	31,672	108,34		8,0820	9,5912
32	7.10.2013	5,6355	415,53	5,8009	662,78	31,854	108,95		8,1270	9,6326
33	4.10.2013	5,6315	414,89	5,7787	661,27	31,754	108,78		8,1145	9,5814
34	3.10.2013	5,5924	414,61	5,7703	660,56	31,728	108,70		8,1090	9,6605
35	2.10.2013	5,6246	415,79	5,8136	664,32	31,758	109,03		8,1320	9,7494

Figure 24- Historical Exchange rate

5.3 Data preparation

For this Time series forecasting model development, the exchange of Euro will be forecasted using historical data. SAP HANA PAL Time series double exponential smoothing algorithm will be used for this forecasting model. The algorithm is suitable to model the time series with trend but without seasonality (Stephen L. Bernard, 2011), because currency exchange is traded daily and doesn't depend on any season. The data that the forecasting model needs is required in a specific table format. This requires that the input data for the model is adjusted to meet the forecasting algorithm requirement. Also, the algorithm doesn't allow empty rows or any null values which means that all empty rows will have to be filled or removed from the data set.

The data preparation includes the following

- Removing empty rows from the dataset.
- Removing unwanted columns.
- Inclusion of additional ID Column in the file

The extracted file has been converted to the required format by the HANA PAL algorithm for example additional column has been added and this file will be transferred to the database table.

Figure 24 below shows the converted file with additional ID column.

ID	Date	Exchange Rate
1	4.1.1999	8,8550
2	5.1.1999	8,7745
3	6.1.1999	8,7335
4	7.1.1999	8,6295
5	8.1.1999	8,5900
6	11.1.1999	8,5585
7	12.1.1999	8,6100
8	13.1.1999	8,7470
9	14.1.1999	8,7245
10	15.1.1999	8,7150
11	18.1.1999	8,6575
12	19.1.1999	8,6300
13	20.1.1999	8,6000
14	21.1.1999	8,6050
15	22.1.1999	8,6225
16	25.1.1999	8,6125
17	26.1.1999	8,6125
18	27.1.1999	8,5985
19	28.1.1999	8,5700
20	29.1.1999	8,5785
21	1.2.1999	8,5395
22	2.2.1999	8,5845
23	3.2.1999	8,6250
24	4.2.1999	8,6425
25	5.2.1999	8,6725

Figure 25 - Converted Historical Exchange rate table

The figure below requires 2 columns as required by the algorithm (SAP, 2013) The ID column represent the key field of individual record and this will be created manually. The raw data column represents the exchange rate.

Table	Column	Column Data Type	Description
Data	1st column	Integer	ID
	2nd column	Integer or double	Raw data

Figure 26 - Double exponential smoothing Input table (SAP, 2013).

To transfer the data from the excel file into the database, a table without any data is created and this table will have 2 columns, the ID and exchange rate fields. This table has same number of columns and data type as required by the input table of the algorithm (SAP, 2013)

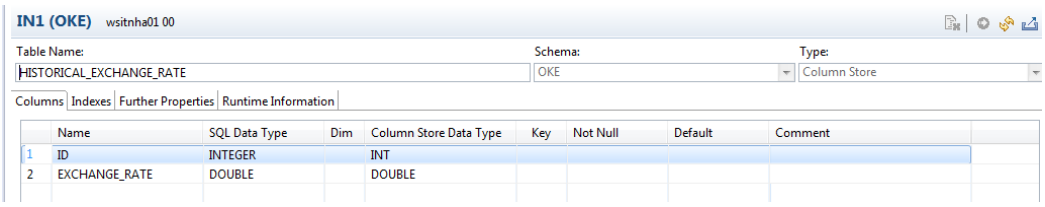

Figure 27 - Double exponential smoothing Input table (SAP, 2013).

At this stage the date field from the extracted file is not required as it is now replaced by the new additional ID column. Figure 27 below shows that the Date field is not mapped to the table in HANA database.

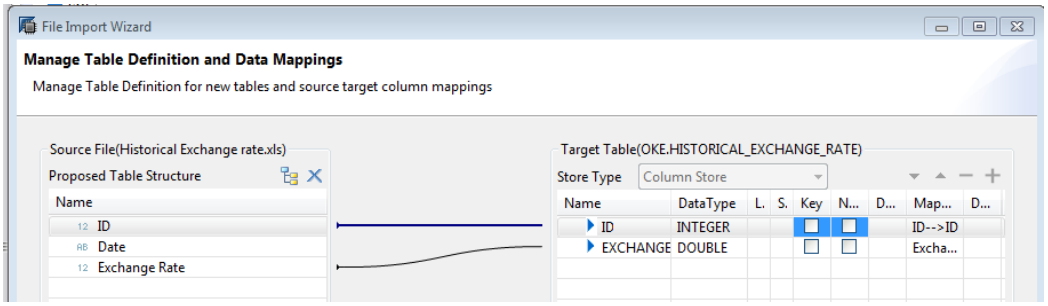

Figure 28 - Data mapping between the file and database table

The data transferred into the table can be seen in figure 28 and this is the table that will be used for the currency forecasting model.

"OKE"."HISTORICAL_EXCHANGE_RATE"

Raw Data | Distinct values | Analysis

Filter pattern

200 rows retrieved - 120 ms

ID	EXCHANGE_RATE
1	8,855
2	8,7745
3	8,7335
4	8,6295
5	8,59
6	8,5585
7	8,61
8	8,747
9	8,7245
10	8,715
11	8,6575
12	8,63
13	8,6
14	8,605
15	8,6225
16	8,6125
17	8,6125
18	8,5985
19	8,57
20	8,5785

Figure 29 - Database table showing historical data

5.4 Data Processing

The data processed for the forecasting model required 3 different time horizons data samples. The first sample data is for 3 months has 66 observations. The second data sample is 1 year and has 250 different observations. The last sample data is the complete data set of about 3741 observations spanning 166 months.

Below table shows the number of observations across different time horizons.

Time Horizon	**Number of Observations**	**Period**
3 months	66	1.08.2013 - 31.10.2013
12 months	250	1.11.2012 - 31.10.2013
166 months	3741	4.01.1999 - 31.10.2013

Table 3 - Different time horizons

After the data cleansing and transformation of the extracted sample data which includes removing rows that do not have any exchange rate and the additional of additional column, the data were imported into the data base.

Table	**Columns**	**Description**
HISTORICAL_EXCHANGE_RATE	2	storage of the complete data sample
HISTORICAL_EXCHANGE_RATE_YEAR	2	storage of 1 year data sample
HISTORICAL_EXCHANGE_RATE_MONTH	2	storage of 3 month data sample

Table 4 - Sample data tables

The forecasting values of these 3 different forecasting models were compared with the actual values and the results analysed.

The number of observations inserted into the database can be seen from Figure 29 – 31 below. All records were inserted based on the number of rows retrieved from the database tables

"OKE"."HISTORICAL_EXCHANGE_RATE"

Raw Data | Distinct values | Analysis

Filter pattern — 3741 rows retrieved - 329 ms

ID	EXCHANGE_RATE
1	8,855
2	8,7745
3	8,7335
4	8,6295
5	8,59
6	8,5585
7	8,61
8	8,747
9	8,7245
10	8,715
11	8,6575
12	8,63
13	8,6
14	8,605
15	8,6225
16	8,6125
17	8,6125
18	8,5985
19	8,57
20	8,5785

Figure 30 - Full data sample

"OKE"."HISTORICAL_EXCHANGE_RATE_MONTH"

Raw Data | Distinct values | Analysis

Filter pattern — **66 rows retrieved - 95 ms**

ID	EXCHANGE_RATE
3 676	7,8345
3 677	7,8655
3 678	7,855
3 679	7,861
3 680	7,8995
3 681	7,884
3 682	7,817
3 683	7,8085
3 684	7,796
3 685	7,813
3 686	7,8435
3 687	7,9015
3 688	7,904
3 689	7,986
3 690	8,0535
3 691	8,1225
3 692	8,094
3 693	8,083
3 694	8,029
3 695	8,059

Figure 31 - Three month data sample

"OKE"."HISTORICAL_EXCHANGE_RATE_YEAR"

Raw Data | Distinct values | Analysis

Filter pattern — 250 rows retrieved - 113 ms

ID	EXCHANGE_RATE
3 492	7,3705
3 493	7,3305
3 494	7,3425
3 495	7,322
3 496	7,3195
3 497	7,302
3 498	7,3015
3 499	7,303
3 500	7,326
3 501	7,322
3 502	7,3595
3 503	7,3695
3 504	7,3715
3 505	7,335
3 506	7,326
3 507	7,325
3 508	7,329
3 509	7,337
3 510	7,3575
3 511	7,3415

Figure 32 - One Year data sample

The actual exchange rate used in the analysis can be seen from Figure 28 below and this table is updated frequently whenever there is new exchange rate.

Raw Data | Distinct values | Analysis

Filter pattern | 26 rows retrieved - 97 ms

ID	Date:	Rate
3 742	01-nov-2013	8,046
3 743	04-nov-2013	8,0165
3 744	05-nov-2013	8,052
3 745	06-nov-2013	8,0505
3 746	07-nov-2013	8,016
3 747	08-nov-2013	8,1755
3 748	11-nov-2013	8,2065
3 749	12-nov-2013	8,2995
3 750	13-nov-2013	8,338
3 751	14-nov-2013	8,331
3 752	15-nov-2013	8,2535
3 753	18-nov-2013	8,2685
3 754	19-nov-2013	8,2285
3 755	20-nov-2013	8,228
3 756	21-nov-2013	8,2065
3 757	22-nov-2013	8,2065
3 758	25-nov-2013	8,2755
3 759	26-nov-2013	8,2685
3 760	27-nov-2013	8,253
3 761	28-nov-2013	8,276
3 762	29-nov-2013	8,32
3 763	02-des-2013	8,3095
3 764	03-des-2013	8,29
3 765	04-des-2013	8,3105
3 766	05-des-2013	8,4035
3 767	06-des-2013	8,434

Figure 33 - Actual exchange rates

5.5 Data analysis and results

The analysis of the result started by comparing the different time horizons forecasted exchange rate values against actual exchange rates. The comparison can be seen in figure 26 where the different time horizon sample data can be measured against the actual value and the model accuracy ascertained.

The model development is a combination of 4 different tables, three of the forecasting tables were automatically generated by the time series double exponential forecasting model and the fourth table holds the data of the actual currency exchange rate. The actual exchange rate started from 1st of November and this table will be updated regularly with new actual exchange rate to measure the accuracy of the model and compare the actual exchange rate against the forecasted exchange rate.

During the analysis of the model and comparison of the forecasted results, the actual exchange rate extracted from the website and updated into the database was joined with the forecasted model generated table into a new table. The new table is a view of the four different tables and contains no data of its own.

It can be seen in figure 33 that the data sample of 3 months produces the smallest variance when compared with the actual exchange rate. The 1 year data sample used in the model showed the biggest variance when compared with the actual data. The full data sample shows a result that is considerably higher than the forecasted value.

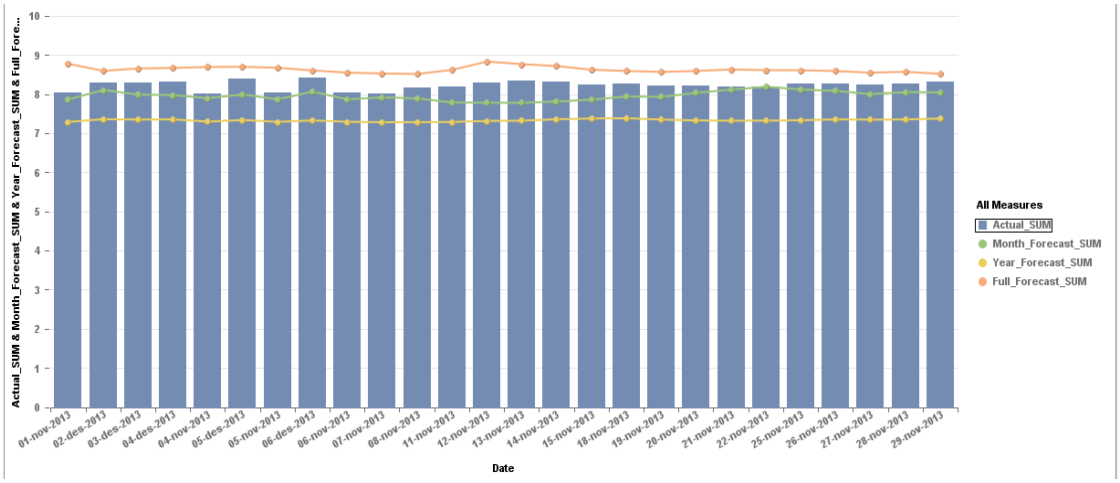

Figure 34 - Actual versus forecasted result Chart

During the data examination, the reason why these variances occur can be attributed to how the double exponential smoothing forecast values. In double exponential smoothing, more weights are given to the latest exchange rates values compared to smaller weights given to older exchange rates. The 1 year data sample from 1st of November 2012 to 31st of October 2013 exchange rates is when the Norwegian Krone

performed stronger than the Euro, and this related to news about the breakup of the Euro currency and the political and economic situations in Greece and Spain due to the debt crisis (Ben Rooney, 2012).

Chart | Table | Grid | HTML

Date	Actual_SUM	Month_Forecast_SUM	Year_Forecast_SUM	Full_Forecast_SUM
01-nov-2013	8,046	7,866	7,29	8,774
02-des-2013	8,31	8,107	7,359	8,595
03-des-2013	8,29	7,987	7,349	8,658
04-des-2013	8,31	7,975	7,355	8,67
04-nov-2013	8,016	7,896	7,298	8,694
05-des-2013	8,404	7,991	7,338	8,7
05-nov-2013	8,052	7,868	7,291	8,674
06-des-2013	8,434	8,064	7,328	8,605
06-nov-2013	8,05	7,867	7,293	8,55
07-nov-2013	8,016	7,919	7,283	8,527
08-nov-2013	8,176	7,891	7,284	8,516
11-nov-2013	8,206	7,788	7,289	8,621
12-nov-2013	8,3	7,781	7,314	8,834
13-nov-2013	8,338	7,777	7,321	8,766
14-nov-2013	8,331	7,811	7,357	8,719
15-nov-2013	8,254	7,86	7,378	8,624
18-nov-2013	8,268	7,939	7,384	8,593
19-nov-2013	8,228	7,93	7,35	8,568
20-nov-2013	8,228	8,033	7,328	8,593
21-nov-2013	8,206	8,115	7,321	8,63
22-nov-2013	8,206	8,189	7,323	8,614
25-nov-2013	8,276	8,12	7,332	8,611
26-nov-2013	8,268	8,083	7,354	8,59
27-nov-2013	8,253	8,002	7,347	8,55
28-nov-2013	8,276	8,051	7,352	8,571
29-nov-2013	8,32	8,046	7,376	8,518

Figure 35 - Actual versus forecasted result table

Table 5 below shows the different statistical methods used in evaluating performance. 3 months data sample has the least RMSE and variance while the 1 year data sample has the highest RMSE and variance. Overall the model is considered a good model. The calculations for the statistical methods can be seen in the appendix section.

	RMSE	MAE	VARIANCE
3 months data sample	0.30	0.27	7.1
1 year data sample	0.90	0.9	23.4
Full data sample	0.41	-0.39	-10.0

Table 5 - Statistic methods of model accuracy

6 Conclusion and future research

This is the final chapter of the thesis and it covers discussion related to conclusion and future research.

6.1 Conclusion

The thesis paper began by reviewing the different disciplines related to predictive analytics such as business intelligence, data mining, data warehouse and predictive analytics.

The objective of this thesis is to develop a currency exchange rate predictive model using time series algorithm and measure the accuracy of the model by comparing the predicted exchange rate against the actual exchange rate. The historical data collected for the research was extracted from Norwegian bank online database and has 3741 transactional records from the period of 4 January 1999 to 31 October 2013. Two other data samples namely, 12 months and 3 months of data were constructed from the original data sets. The extracted data doesn't contain exchange rates from weekends and national public holidays.

Three different predictive models for each data sample were developed and measured against the actual result and based on the analysis of the result, it can be concluded that the most recent data which is the 3 months data sample produces the least variance when compared to the actual result and gives better accuracy. However, the model did not outperform the random walk theory which stated that future currency exchange rates cannot be predicted based on historical and current data because exchange rates follow an unpredictable path (Xie, Xin, 2011: 25)

6.2 Future research

Technical analysis of currency exchange rate is not adequate enough to give good accuracy as it only uses quantitative data based on historical and current data to predict future exchange rate. What technical analysis lacks is the information related to the driving forces behind the rise and fall on currency exchange rates for example an oil exporting country such as Norway that is heavily dependent on the price of oil may experience high exchange rate appreciation when the oil prices rise and depreciation when it falls (Quaisar F. Akram,2012). Also political and economic situation such as interest rate, debt crisis can also affect currency exchange.

Previous Studies using news headlines (Desh P. and Raymond .W, 2001) to predict exchange have been conducted and it will be worthwhile to use text analysis to quantify fundamental analysis such as market sentiments and various economic factors into numbers and integrate it with technical analysis to get better accuracy.

7 References

Mike Gualtieri (2013), The Forrester wave: Big data predictive analytics solutions Q1 2013.

Galit Shmueil and Otto R. koppius (2011:9), The Predictive Analytics in information systems research.

Big Data for customer experiences, http://adayinbigdata.com [Accessed on 18.08.2013]

Gartner hype cycle 2013. Cycle for Emerging Technologies Maps out Evolving Relationship between Humans and machines, (2013). http://www.gartner.com/newsroom/id/2575515 [Accessed on 07.09.2013]

M. Zaman, (2013), Predictive Analytics: the Future of Business Intelligence: http://www.studymode.com/essays/Predictive-Analytics-The-Future-Of-Business-707768.html [Accessed on 18.09.2013]

Zhenyu Huang, Lei-da Chen, and Mark N. Frolick (2002) INTEGRATING WEB-BASED DATA INTO A DATA WAREHOUSE, 23

Tim Chenoweth, Karen Corral, and Haluk Demirkan (2006) Communications of the ACM, 115

Searchdatamanagement (2006) http://searchdatamanagement.techtarget.com/definition/business-intelligence [Accessed on 18.09.2013]

Surajit Chaudhuri, Umeshwar Dayal, and Vivek Narasayy (2011) An Overview of Business Intelligence Technology, p.90, VOL 58, Art. # 8

Yellowfin International 2010, making business intelligence easy, White paper in-memory analytics, p.3

Shiraj Khan, Auroop R Ganguly and Amar Gupta, Data Mining and Data fusion for Enhanced Decision Support (2008), Springer Berlin Heidelberg

Jiawan Han and Micheline Kamber Data Mining concepts and Techniques 2nd edition (2006), p.6, Burlington, MA, USA, Elsevier Science & Technology

Land Collete, The Power of Predictive Analytics, 2006, p.12

Wayne W. Eckerson,(2007) Predictive Analytics Extending the Value of Your Data Warehousing Investment p.5, SAS Institute.

Delloite, In memory computing, the holy grail of analytics,(2013), p.4, Delloite and Touche

Linear regression: http://docentes.deio.fc.ul.pt/kfturkman/regression.pdf [Accessed on 20.11.2013]

Alan Sykes, An introduction to regression analysis, p.1.

Model Extremely Complex Functions, Neural Networks: http://www.statsoft.com/textbook/neural-networks/ [Accessed on 16.10.2013]

Berry de Ville (2006), Decision Trees for Business Intelligence and Data Mining: Using SAS Enterprise Miner, Cary, NC, USA, SAS Institute Inc

Michael Berry and Gordon Linoff, Data Mining Techniques: For Marketing, Sales and Customer Relationship Management, 2nd Edition, 2004, Indianapolis, Indiana, Wiley publishing Inc

Prof. Erik Mooi and Prof. Marko Sarstedt,(2011), A concise guide to Market Research, The process, Data and Methods using IBM SPSS Statistics, Springer Books.

Alejandro Correa, Andres Gonzalez, Catherine Nieto and Darwin Amezquita, Constructing a Credit Risk Scorecard using Predictive Cluster, SAS Global Forum 2012.

http://pegasus.cc.ucf.edu/~cwang/sta6714/Lecture8/Notes/Association%20Analysis.pdf [Accessed on 20.11.2013]

Sanford Gayle (2000), SAS Institute, the Marriage of Market Basket Analysis to Predictive Modeling

Beniwal and Arora (2012), classification and feature selection techniques in data mining, VOL. 1 Issue 6.

Christian Mihaesu (2002),Naive-Bayes Classification Algorithm,p.1,http://software.ucv.ro/~cmihaescu/ro/teaching/AIR/docs/Lab4-NaiveBayes.pdf [Accessed on 20.11.2013]

http://www.princeton.edu/~achaney/tmve/wiki100k/docs/Naive_Bayes_classifier.html [Accessed on 20.11.2013]

RC Chakraborty (2010), Fundamentals of Genetic Algorithms http://www.myreaders.info/09_Genetic_Algorithms.pdf [Accessed on 20.11.2013]

Cohen, Barry H. (2012), Introductory Statistics for the Behavioral Sciences (7th Edition), p.459

http://blog.minitab.com/blog/adventures-in-statistics/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit [Accessed on 20.11.2013]

Saravanan Thirumuruganathan (2007), A detailed introduction to K-Nearest Neighbor (KNN) Algorithm, http://saravananthirumuruganathan.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighbor-knn-algorithm/[Accessed on 02.01.2014]

Scott Fortmann-Roe (2012), Understanding the Bias-Variance Tradeoff, http://scott.fortmann-roe.com/docs/BiasVariance.html[Accessed on 20.11.2013

http://wps.ablongman.com/wps/media/objects/2829/2897573/ch18.pdf [Accessed on 20.11.2013]

http://www.norges-bank.no/en/prisstabilitet/valutakurser/ [Accessed on 20.11.2013]

SAP HANA Predictive Analysis Library (PAL), (2013)

http://online.wsj.com/news/articles/SB10001424052702304665904576384111852016334[Accessed on 18.11.2013]

http://www.it.iitb.ac.in/~praj/acads/seminar/04329008_ExponentialSmoothing.pdf [Accessed on 20.11.2013]

http://www.statoek.wiso.unigoettingen.de/veranstaltungen/graduateseminar/SmoothingMethods_Narodzonek-Karpowska.pdf[Accessed on 20.11.2013]
http://www.sas.com/events/cm/174390/assets/102892_0107.pdf [Accessed on 20.11.2013]

Padraic .G Neville (1999), Decision Trees for Predictive Modeling, p.2,SAS Institute Inc.

G Kirchgassner and J. Wolters,(2012), Introduction to Modern Time Series,p.2, Berlin, Springer books

Lamonth J, Ph.D, Predictive analytics: an asset to retail banking worldwide, 2005 VOL 14, Issue 10, http://www.kmworld.com/Articles/Editorial/Features/Predictive-analytics-an-asset-to-retail-banking-worldwide-14587.aspx [Accessed on 04.12.2013]

http://events.asug.com/2012BOUC/0805_Demonstration_of_SAP_Predictive_Analysis_1_0_consumption_from_SAP_BI_clients_and_best_practices.pdf [Accessed on 04.12.2013]

Ben Rooney (2012), Europe's debt crisis: 'No clear end in sight'.
http://money.cnn.com/2012/01/06/markets/europe_debt_crisis/ [Accessed on 05.12.2013]
Zhenyu Huang, Lei-da Chen, and Mark N. Frolick (2002) INTEGRATING WEB-BASED DATA INTO A DATA WAREHOUSE, 23

Xie, Xin (2011) Full-View Integrated Technical Analysis: A Systematic Approach to Active Stock Market Investing, p.25, John Wiley & Sons Pte (Asia) Ltd.

http://www.econometricsociety.org/meetings/esem02/cdrom/papers/1223/QFAkram.pdf [Accessed on 07.12.2013]

http://crpit.com/confpapers/CRPITV5Peramunetilleke [Accessed on 07.12.2013]

http://www.saphana.com/docs/DOC-4142 [Accessed on 09.12.2013]

http://www.saedsayad.com/model_evaluation_r.htm [Accessed on 09.12.2013]

Saed Sayad, (2013), Artificial Neural Network, http://www.saedsayad.com/artificial_neural_network.htm

http://news.contextuall.com/what-is-the-mean-absolute-error/[Accessed on 09.12.2013]

Dunham, M. H. (2003), Data mining introductory and advanced topics. Upper Saddle River, NJ: Pearson Education, Inc.

Rebecca Bundhun, Cost of summer getaways hit as air ticket rise (2012), http://www.thenational.ae/business/industry-insights/tourism/cost-of-summer-getaways-hit-as-air-ticket-prices-rise [Accessed on 27.12.2013]

Stephen Link, War and its effect on oil prices (2011), http://voices.yahoo.com/war-its-effect-oil-prices-10155733.html [Accessed on 27.12.2013]

Statsofta (2013), http://statisticasoftware.wordpress.com/tag/data-mining/ [Accessed on 30.12.2013]

Reshma H. Bonde, Shital D.Tatale, Rashmi V. Sawalakhe and Prashant C. Jikar, (2013), Neural Network and Fuzzy logic, International Journal of Research in Engineering and Technology (IJRET), VOL. 2 Issue 5.

Fan-Yin Tzeng and Kwan-Liu Ma, (2005), Opening the black box – Data Driven Visualization of Neural Networks.

Carlos Gershenson , (2003), Artificial Networks for Beginners

Hery Mulyana (2011), Time Series with SPSS, http://www.scribd.com/doc/58449449/Time-Series-With-SPSS[Accessed on 27.12.2013]

Larson JS, Bradlow ET, Fader PS (2005) An exploratory look at supermarket shopping paths.Int J Res Mark 22(4):395–414

Jim Frost (2013), Regression Analysis: How Do I Interpret R-squared and Assess the Goodness-of-Fit?, http://blog.minitab.com/blog/adventures-in-statistics/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit [Accessed on 27.12.2013]

Cohen, Barry H., (2012) Introductory Statistics for the Behavioral Sciences (7th Edition), p.459

Wayne W. Eckerson (2007), Predictive analytics, Extending the value of your data warehousing investment,p.5,http://www.sas.com/events/cm/174390/assets/102892_0107.pdf [Accessed on 20.11.2013]

Robert Kunst (2012), Econometric Forecasting, p.8, http://homepage.univie.ac.at/robert.kunst/prognose.pdf[Accessed on 01.01.2014]

Kerby Shedden (2013), Multiple linear regression, p.1,http://dept.stat.lsa.umich.edu/~kshedden/Courses/Stat401/Notes/401-multreg.pdf[Accessed on 01.01.2014]

Ricky Ho (2012), Predictive Analytics: Evaluate Model Performance, http://horicky.blogspot.fi/2012/06/predictive-analytics-evaluate-model.html[Accessed on 01.01.2014]

Fabricio Voznika and Leonardo Viana (2007), Data mining classification, p.1,http://courses.cs.washington.edu/courses/csep521/07wi/prj/leonardo_fabricio.pdf[Accessed on 01.01.2014]

Tom M. Mitchel (2006), The Discipline of Machine Learning, p.4, http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf [Accessed on 02.01.2014]

Intel IT Center (2013), http://www.intel.com/content/dam/www/public/us/en/documents/best-practices/big-data-predictive-analytics-overview.pdf [Accessed on 05.01.2014]

Marko Grobelnik (2012), http://www.planet-data.eu/sites/default/files/presentations/Big_Data_Tutorial_part4.pdf [Accessed on 05.01.2014]

Quaisar F. Akram (2012), http://www.econometricsociety.org/meetings/esem02/cdrom/papers/1223/QFAkram.pdf [Accessed on 07.12.2013]

Desh P. ,Raymond .W, (2001 http://crpit.com/confpapers/CRPITV5Peramunetilleke [Accessed on 07.12.2013]

Appendices

Appendix 1. Actual exchange rate sql script

```
CREATE COLUMN TABLE "OKE"."ACTUAL_EXCHANGE_RATE" ("ID" INTEGER
CS_INT NOT NULL ,
            "Date:" NVARCHAR(11),
            "Rate" DOUBLE CS_DOUBLE,

            PRIMARY KEY ("ID")) UNLOAD PRIORITY 5 AUTO MERGE
```

Table 6 - Actual exchange rate sql script

Appendix 2. 166 months exchange rate sql script

```
CREATE COLUMN TABLE "OKE"."HISTORICAL_EXCHANGE_RATE" ("ID" INTE-
GER CS_INT,

            "EXCHANGE_RATE" DOUBLE CS_DOUBLE) UNLOAD PRIORITY 5
AUTO MERGE
```

Table 7 – 166 months exchange rate sql script

Appendix 3. 3 months exchange rate table

```
CREATE COLUMN TABLE "OKE"."HISTORICAL_EXCHANGE_RATE_MONTH" ("ID" INTEGER
CS_INT,

            "EXCHANGE_RATE" DOUBLE CS_DOUBLE) UNLOAD PRIORITY 5 AUTO MERGE
```

Table 8 - 3 months exchange rate table

Appendix 4. 1 Year data sample table sql script

```
CREATE COLUMN TABLE "OKE"."HISTORICAL_EXCHANGE_RATE_YEAR" ("ID"
INTEGER CS_INT,

            "EXCHANGE_RATE" DOUBLE CS_DOUBLE) UNLOAD PRIORITY 5
AUTO MERGE
```

Table 9 - 1 Year data sample table sql script

Appendix 5. Exchange rate comparison table

```
CREATE VIEW "OKE"."EXCHANGE_RATE_COMPARISM_TABLE" ( "Date",
            "Actual",
            "Month_Forecast",
            "Year_Forecast",
            "Full_Forecast" ) AS select
            T0."Date:" "Date",
            T0."Rate" "Actual",
            T2."Month_Forecast" "Month_Forecast",
            T3."Year_forecast" "Year_Forecast",
            T1."Full_forecast" "Full_Forecast"
from "OKE"."ACTUAL_EXCHANGE_RATE" T0
left outer join
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Result" T3 on T0."ID" = T3."ID"
left outer join
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Result" T2 on T0."ID" = T2."ID"

left                                outer                                join
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Result" T1 on T0."ID" = T1."ID" WITH READ ONLY
```

Table 10 - Exchange rate comparison table

Appendix 6. Full data sample double smoothing Algorithm

```
DROP TABLE
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Args";
CREATE COLUMN TABLE
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Args" ("NAME" VARCHAR(50), "INTARGS" INTEGER, "DOUBLEARGS"
DOUBLE, "STRINGARGS" VARCHAR(100));
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Args" VALUES ('RAW_DATA_COL', 1, null, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Args" VALUES ('ALPHA', null, 0.9, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Args" VALUES ('BETA', null, 0.7, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Args" VALUES ('STARTTIME', 3741, null, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Args" VALUES ('FORECAST_NUM', 30, null, null);
DROP TABLE
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Result";
CREATE COLUMN TABLE
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Result" ("ID" INTEGER, "Full_forecast" DOUBLE);
CALL
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el"("OKE"."HISTORICAL_EXCHANGE_RATE",
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Args",
"OKE"."Currency_Forecast_Predictive_Model::Full_forecast_rate.All_Periods_Mod
el_doubleSmooth_Result") with overview;
```

Table 11 - Full data sample double smoothing Algorithm

Appendix 7. 3 months data sample double smoothing Algorithm

```
DROP TABLE
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Args";
CREATE COLUMN TABLE
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Args" ("NAME" VARCHAR(50), "INTARGS" INTEGER, "DOUBLEARGS"
DOUBLE, "STRINGARGS" VARCHAR(100));
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Args" VALUES ('RAW_DATA_COL', 1, null, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Args" VALUES ('ALPHA', null, 0.9, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Args" VALUES ('BETA', null, 0.6, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Args" VALUES ('STARTTIME', 3741, null, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Args" VALUES ('FORECAST_NUM', 30, null, null);
DROP TABLE
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Result";
CREATE COLUMN TABLE
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Result" ("ID" INTEGER, "Month_Forecast" DOUBLE);
CALL
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del"("OKE"."HISTORICAL_EXCHANGE_RATE_MONTH",
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Args",
"OKE"."Currency_Forecast_Predictive_Model::Months_forecast.Months_forecast_mo
del_doubleSmooth_Result") with overview;
```

Table 12 - 3 months data sample double smoothing Algorithm

Appendix 8. 1 Year data sample double smoothing Algorithm

```
DROP TABLE
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Args";
CREATE COLUMN TABLE
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Args" ("NAME" VARCHAR(50), "INTARGS" INTEGER, "DOUBLEARGS" DOU-
BLE, "STRINGARGS" VARCHAR(100));
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Args" VALUES ('RAW_DATA_COL', 1, null, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Args" VALUES ('ALPHA', null, 0.7, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Args" VALUES ('BETA', null, 0.3, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Args" VALUES ('STARTTIME', 3740, null, null);
INSERT INTO
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Args" VALUES ('FORECAST_NUM', 30, null, null);
DROP TABLE
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Result";
CREATE COLUMN TABLE
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Result" ("ID" INTEGER, "Year_forecast" DOUBLE);
CALL
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model"
("OKE"."HISTORICAL_EXCHANGE_RATE_YEAR",
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Args",
"OKE"."Currency_Forecast_Predictive_Model::Year_forecast.Year_forecast_model_
doubleSmooth_Result") with overview;
```

Table 13 - 1 Year data sample double smoothing Algorithm

Appendix 9. 3 months data sample RMSE

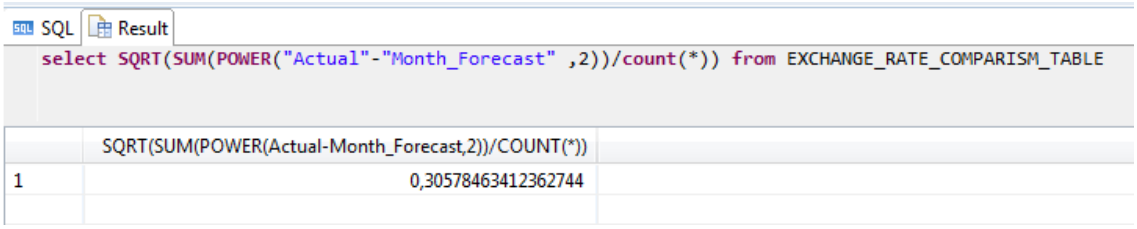

Figure 36 - 3 months data sample RMSE

Appendix 10. 1 year data sample RMSE

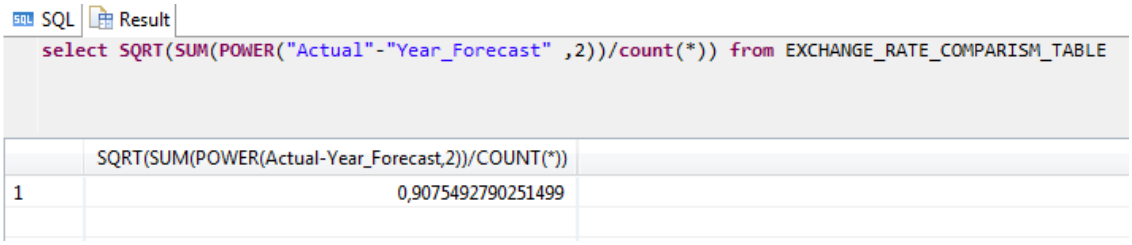

Figure 37 - 1 year data sample RMSE

Appendix 11. Full data sample RMSE

SQL Result

```
select SQRT(SUM(POWER("Actual"-"Full_Forecast" ,2))/count(*)) from EXCHANGE_RATE_COMPARISM_TABLE
```

	SQRT(SUM(POWER(Actual-Full_Forecast,2))/COUNT(*))
1	0,41719585214864147

Figure 38 - Full data sample RMSE

Appendix 12. 3 months data sample variance

SQL Result

```
select SUM("Actual"-"Month_Forecast") from OKE.EXCHANGE_RATE_COMPARISM_TABLE
```

	SUM(Actual-Month_Forecast)
1	7,104921446316573

Figure 39 - 3 months data sample variance

Appendix 13. 1 year data sample variance

SQL Result

```
select SUM("Actual"-"Year_Forecast")  from OKE.EXCHANGE_RATE_COMPARISM_TABLE
```

	SUM(Actual-Year_Forecast)
1	23,468507828080305

Figure 40 - 1 year data sample variance

Appendix 14. Full data sample variance

SQL Result

```
select SUM("Actual"-"Full_Forecast")  from OKE.EXCHANGE_RATE_COMPARISM_TABLE
```

	SUM(Actual-Full_Forecast)
1	-10,303553552866976

Figure 41 - Full data sample variance

Appendix 15. 3 months data sample MAE

SQL Result

```
select SUM("Actual"-"Month_Forecast") / COUNT(*) from OKE.EXCHANGE_RATE_COMPARISM_TABLE
```

	SUM(Actual-Month_Forecast)/COUNT(*)
1	0,2732662094737143

Figure 42 - 3 months data sample MAE

Appendix 16. 1 year data sample MAE

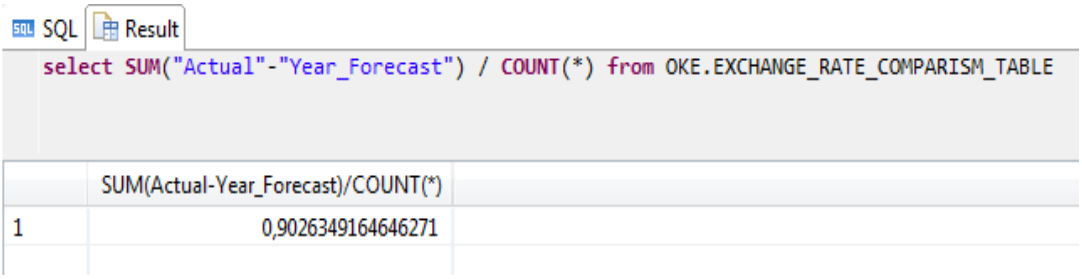

Figure 43 - 1 year data sample MAE

Appendix 17. Full month data sample MAE

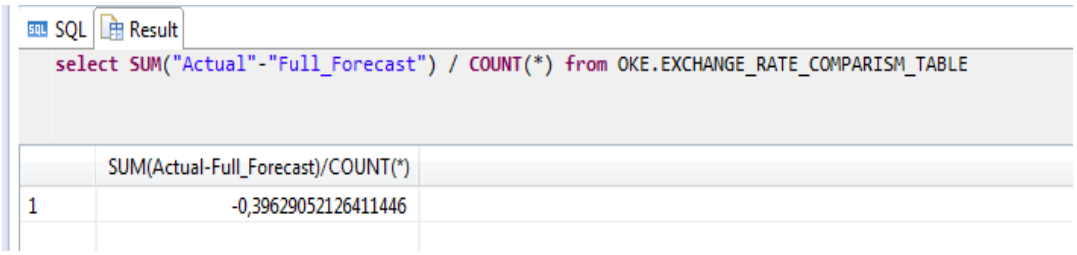

Figure 44 - Full month data sample MAE