ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

COLLEGE OF NATURAL SCIENCES

DEPARTMENT OF COMPUTER SCIENCE

Predictive Model for ECX Coffee Contracts

Frehiwot Mulugeta Deressa

A THESIS SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES OF ADDIS ABABA UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN COMPUTER SCIENCE

October 2014

Addis Ababa

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

COLLEGE OF NATURAL SCIENCES

DEPARTMENT OF COMPUTER SCIENCE

Predictive Model for ECX Coffee Contracts

Frehiwot Mulugeta Deressa

Advisor: Sebsibe Hailemariam(Ph.D)

Approved by Board of Examiners:

Name	Signature
1. Sebisbe Hailemariam (Ph.D), Advisor	___________________
2. Yaregal Assabie(PhD), Examiner	___________________
3. Dida Midekso (PhD), Examiner	____________________

Acknowledgement

First and foremost, I would like to express my deepest gratitude to my advisor Dr. Sebsibe Hailemariam for his continuous support and professional guidance throughout this thesis work. I would like to extend my thanks to staff of ECX especially to market data specialists for their cooperation. My deepest gratitude goes to my family who were always motivating me to work hard. I would also like to thank my friends for their support and companionship.

Table of Content

List of Figures

List of Tables

Acronyms

ANN	Artificial Neural Network
BP	Backward Propagation
CRISP-DM	Cross-Industry Standard Process for Data Mining
ECX	Ethiopia Commodity Exchange
FFNN	Feed Forward Neural Network
KDD	Knowledge Discovery in Databases
NYBOT	New York board trade
ML	Machine learning
MLP	Multilayer Preceptor
MSE	Mean Squared Error
R^2	Coefficient of determination
RBF	Radial Basis Function
SEMMA	Sample, Explore, Modify, Model, Assess (data mining)

Abstract

Ethiopia Commodity Exchange is a commodity market that transforms the traditional agricultural marketing system into modern and transparent market. Ethiopia is known for its high quality and highly diversified type of coffee and ECX has designed detailed coffee contract and the market executes many trades for these contracts. This research aims to study the relationship between ECX coffee contract and to propose prediction model that assists the market to undertake efficient coffee trading system. The price prediction model will be used to predict the daily selling price of all coffee contracts. The prediction model was developed by the most widely used machine learning method, Artificial Neural Network.

Five and half years of ECX coffee trading data have been used to analyze the problem, to train and test the models. The coffee trading data have been studied intensively by correlation coefficient and scatter plot matrixes using volume of coffee traded in the market and availability of the contract in a year. It was found that washed Sidama coffee A grade 3 (WSDA3) contract was traded in a larger volume and available throughout the year. Moreover, the contract is highly correlated with most coffee contracts. And thus WSDA3 was selected as a reference contract to represent all export coffee contracts.

Coffee contracts daily price data show non-linear characteristics. Traditional statistical methods are unable to develop prediction model for non-linear data. Artificial neural network can flexibly model linear or non-linear relationship between variables. Among the artificial network algorithms; the radial basis function neural network (RBF) and multilayer perceptron neural network (MLP) are used to approximate any linear or non-linear function. MLP and RBF methods were employed to develop coffee contracts price prediction model.
Three experiments were designed to build the coffee contract price prediction models. For washed Sidama coffee, for unwashed Sidama coffee contracts and for contracts different from Sidama origin. The performance of the models was evaluated on the test data set by coefficient of determination and mean squared error.

The experimental result reveals that large R^2 values with small variance were obtained in MLP based models than RBF Based models. Moreover, the smallest MSE with small variance is observed in MLP based models as compared to models constructed by RBF algorithm. The results obtained from the study showed that the MLP networks are capable of predicting the daily price of coffee contracts than the RBF networks because MLP networks are global function approximators.

In MLP base models, the largest R^2 with smallest variance is achieved in Sidama washed coffee and different origin washed coffee contracts. Similarly, MLP based models the smallest MSE with minimum variance is achieved Sidama washed coffee and different origin washed coffee. Sidama washed coffee and different origin washed coffee contracts respectively. The accuracy results of washed coffee contracts using MLP algorithms are higher than unwashed contracts. Generally, coffee contract that belong to the same processing type to the reference contract (WSDA3) has higher accuracy result than that of contract in different processing type.

Key words: ANN, Coffee Contract, ECX, MLP, Machine Learning, RBF, Price Prediction

CHAPTER ONE
INTRODUCTION

The Ethiopia Commodity Exchange (ECX) establishment was founded in April 2007, Proclamation number 550/2007, with a vision to revolutionize Ethiopian traditional agricultural marketing system through creating a new market place where market actors, such as farmers, traders, processors, exporters and consumers come together to trade a commodity assured of quality, quantity, payment and delivery. Its mission is to connect all buyers and sellers in an efficient, reliable and transparent market by harnessing innovation and technology [13]. The fact of bringing together the buyers and sellers at a central market place at any point in time results a greatest concentration of traders that allow maximum effective competition among buyers and sellers [19].

According to Article 2 No.6 of the ECX proclamation No. 550/2007, commodity exchange means an exchange where standardized commodity[1] contracts are traded. A contract is a specification for a commodity which is traded in the exchange market. A contract is defined as a standard agreement to buy or sell a specific commodity detailing the amount, grade, price of a commodity and delivery place. Since ECX is a spot market [2] delivery and payment for the commodity is immediate.

As indicated by [20] ECX commenced trading operations in April 2008 by trading maize contracts. Currently, Ethiopian Commodity Exchange offers standardized trading contracts for maize, wheat, coffee, haricot beans, and sesame,

Globally, there are two main families of coffee seeds: Arabica and Robusta based on their growths and qualities. Arabica coffee beans, which grow mainly in the tropical highlands of the Western Hemisphere, make up the bulk of world production. While Robusta coffee beans are produced largely in the low, hot areas of Africa and Asia. Arabica can be found in several

---

[1] Commodity is a product having commercial value that can be produced, bought and sold

[2] Ethiopia Commodity Exchange Proclamation number 550/2007; Article 2 No.15 defines Spot Market as an exchange trade contract for immediate delivery and payment of the commodity.

varieties with unique flavors and it is generally considered the high-quality coffee. On the other hand, Robusta has high caffeine content and a more standardized flavor.

Coffee plants are grown in Ethiopia indigenously for the past 3,000 years with many variety and flavor profile [20]. The country is known to be the birth place for coffee and it is recognized all over the world for its high quality coffee. According to Dominic and Surendra [11] there are different varieties of Arabica coffee that are grown in different regions of Ethiopia. Coffee in different regions has got distinct bean characteristics due to their geographical location and climatic differences.

There is no universally accepted grading and classification system for coffee. Each producing country has developed its own classification and grade charts. Some of the commonly used criteria are origin, preparation method, bean size, bean shape, color, number of defects, and cup quality [7].

ECX develops and maintains commodity grades as a standard for the purpose of trading the commodities on the market. In the ECX market, coffee is classified based on its origin, class, the preparation type, grade of coffee and market type [20]. First, coffee is classified based on its origins; origin is the broader area classification of coffee. Currently, the market has 22 coffee origins; they are Anderacha, Bale, Bebeka, Bench-Maji, Bi-Product, East Wollega, Forest, Gelana-Abaya, Gimbi, Godere, Harar, Jimma, Keffa, Kelem Wollega, Kochere, Lekemite, Limmu, Tepi, Sidama, Wenago, Yeki and Yirgachefe. These coffee origins are classified by their class or locality that is the specific woredas where the coffee is produced. Each origin coffee has 1 to 5 class(es) represented by A, B, C, D and E. Considering Sidama origin coffee, it is classified in to three classes "Sidama A" read as Sidama class A coffee grown in Borena, Benssa, Guji, Arbigona, Bale and West Arsi while Sidama coffee from Aleta Wendo, Dale, Chiko, Dara, Shebedino, Wensho and Loko Abaya are grouped into "Sidama B" and "Sidama C" coffee come from Kembata and Timbaro, Wellayta, South Omo and Gamugoffa [14, 15].

Coffees are further classified by their processing type. Coffee processing is a method of converting the raw fruit coffee into commodity green coffee. There are two methods of coffee processing; washed and unwashed coffee. Washed coffees are green coffee prepared by wet processing of the fruit. On the other hand, unwashed coffee is green coffee prepared by dry processing of the fruit [7].

The coffee that is classified based processing is again classified based on its grade. According to the Ethiopia quality inspection, coffee grade is determined based on green analysis (visual test) and liquor analysis (cup test). The green analysis is made by visual inspection of defects, shape, color and odor which has 40% of the total weight and cup test (acidity, body, and flavor) has 60% weight. Grades are ranked from grade one up to ten and under grade (UG)[3], coffee with the least grade standard.

And finally coffees are classified by their market type; as Specialty, Export and Local coffee. Specialty coffee is coffee that is distinctive because of its full cup taste and little to no defects and that may command a market premium due to its high quality. Washed and unwashed coffee from grade 3 to 9 and UG is categorized under Export coffee. Local coffee is very inferior in quality due to high presence of impurities or if the coffee is stored for a long period and loses its flavor. Currently, ECX accommodates about 421 detailed coffee contracts. This coffee classification is unique to ECX [15].

Table 1.1: Washed export coffee contract

Origin	Coffee Contract	Specific Woreda or Zone	Symbol	Grades	Contracts	Warehouse
YIRGACHEFE	YIRGACHEFE A	Yirgachefe, Wenago, Kochere and Gelana Abaya	WYCA	3 TO 9, UG(p), UG(np)	WYCA3, WYCA4, …	Dilla
	YIRGACHEFE B	Yirgachefe, Wenago, Kochere and Gelana Abaya	WYCB	3 TO 9, UG(p), UG(np)	WYCB3,WYCB4, …	Dilla
SIDAMA	SIDAMA A	Borena (except Gelena/Abaya), Benssa, Guji, Chire, Bona Zuria, Arroressa, Arbigona, Bale Arsi and W. Arsi.	WSDA	3 TO 9, UG(p), UG(np)	WSDA3, WSDA3…	Awassa
	SIDAMA B	Aleta Wendo, Dale, Chiko, Dara, Shebedino, Amaro, Dilla zuria, Wensho and Loko Abaya	WSDB	3 TO 9, UG(p), UG(np)	WSDB3, WSDB3…	Awassa
	SIDAMA C	Kembata and Timbaro, Wellayta, S. Omo and Gamogoffa.	WSDC	3 TO 9, UG(p), UG(np)	WSDC3, WSDC3…	Soddo
LIMMU	LIMMU A	Limmu Seka, Limmu Kossa, Manna, Gomma, Gummay, Seka Chekoressa, Kersa, Shebe and Gera.	WLMA	3 TO 9, UG(p), UG(np)	WLMA3, WLMA4 …	Jimma
	LIMMU B	Bedelle, Noppa, Chorra, Yayo, Alle, and Didu Dedessa.	WLMB	3 TO 9, UG(p), UG(np)	WLMB3, WLMB4…	Bedelle

[3] Under grade (UG) is coffee with lower grade.

KAFFA	KAFFA	Gimbo, Gewata, Chena	WKF	3 TO 9, UG(p), UG(np)	WKF3, WKF4, …	Bonga
TEPI	TEPI	Mezenger (Godere) and Sheka.	WTP	3 TO 9, UG(p), UG(np)	WTP3, WTP4, …	Bonga
BEBEKA	BEBEKA	Bench Maji	WBB	3 TO 9, UG(p), UG(np)	WBB3, WBB4, …	Bonga
LEKEMPTI	LEKEMPTI	Kelem, East and West Wollega.	WLK	3 TO 9, UG(p), UG(np)	WLK3, WLK4, …	Gimbi

Table 1.1 shows washed and unwashed export coffee contracts classification. In Export coffee there are seven washed origins and nine unwashed origins. Each origin may have one or up to five coffee classes. Each coffee class is classified in to 9 grades. Totally ECX have 270 export coffee contracts. The complete ECX export coffee contract is presented in Annex 1 the same types of contracts were designed for specialty, and local coffee contracts [15].

Trading on the ECX shall take place on the trading floor of the ECX located in Addis Ababa, where buyers and sellers or their representative may participate in "open outcry" (shouting) bidding for commodities [14, 49].

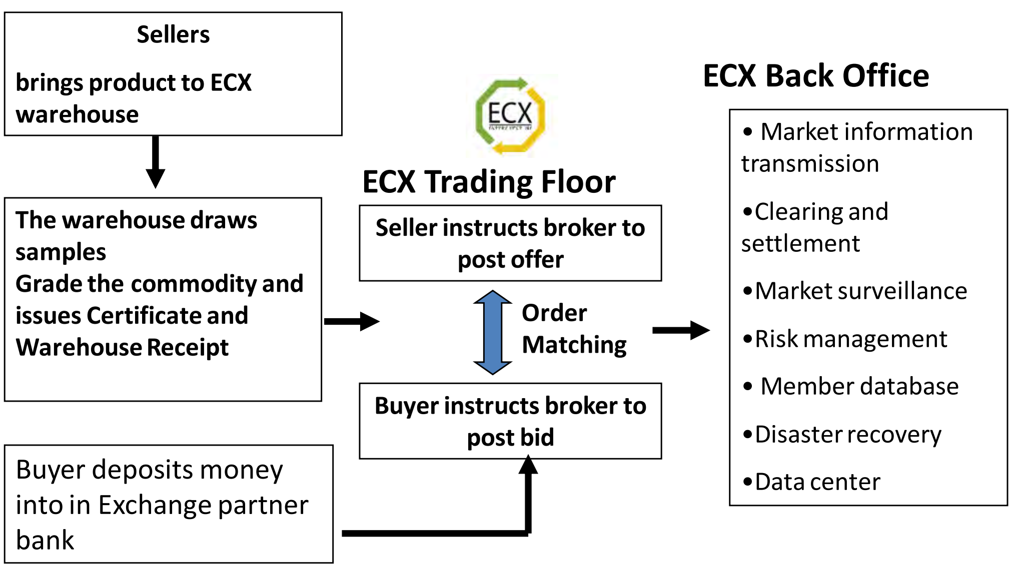


Figure 1.1:- ECX trading system

Figure 1.1 is the pictorial representation of how ECX undertakes its trading activity. As the rule of ECX [49], the seller bring his product or commodity to ECX warehouse for grading while the market take sample from the commodity grade the product and provides electronic warehouse receipt[4] to the seller. On the other hand, commodity buyer should deposit money in the ECX operated banks. Trades are made on the trading floor by bidding or offering a price, grade, quantity of contracts and warehouse; the buyers and sellers meet in person in a trading floor, where they shout out their offers. The ECX automated back office system verifies the availability of buyer funds in a deposit account to cover the payment and that the seller has the sold goods deposited in an ECX warehouse before the trade is accepted by the system. Then an electronic warehouse receipt is sent to the buyer and money is transferred from the buyer's account to the seller's. This automated reconciliation process takes just minutes.

The ECX warehouse is a place where commodities are sampled, weighed, graded and stored. Commodities are deposited in warehouses operated by ECX in major surplus producing regions of the country [49]. Currently, the market has 17 warehouses; out of these nine of them are used for coffee (Table 1.2). A contract might be stored in one or more than one warehouses found in different coffee producing areas of the country.

Table 1.2: ECX coffee warehouses capacity, locations and distance from Addis Ababa

No	Location	Quintals Capacity in tons	Distance From Addis Ababa KM
1	Addis Ababa(Saris)	300,000	-
2	Dire Dawa	50,000	515
3	Hawassa	200,000	273
4	Jimma	170,000	346
5	Bedelle	50,000	580
6	Dilla	280,000	390
7	Gimbi	135,000	450
8	Sodo	110,000	329
9	Bonga	110,000	450

ECX set daily selling price limit for the entire contracts; floor (the minimum price) and ceiling price (the maximum price) and the traders will trade within this range [49]. The daily price limit is calculated as follows: if a contract is sold in the previous day, today's selling price is calculated as plus or minus 5% of last closing price[5] (floor is minus 5% while ceiling is plus 5%). If a contract is not sold in the previous day and if it is sold within the previous two days,

4 Electronic Warehouse Receipt issued by the ECX Central Depository in order to establish legal title to the deposited commodity. It is a receipt that is used to sell commodities in ECX (Rule of ECX, 2010).

5 Last closing price is the average price of contract sold within the last one minute.

today's price limit is calculated as plus or minus 10% of last closing price; else if the contract is not sold in the previous two days but sold within the past 10 days, the limit will be plus or minus 15%. If it is not sold beyond ten days the market would not display the price limit and the trader cannot get price data; they are required to ask the ECX help desk the contracts last selling day closing price or have to search the price information from ECX website but they have to go day by day until the contract's last trade day was found; which requires much time and it is tiresome.

Currently, the market eliminates the price limit due to the coffee price fluctuation in the international coffee market. And thus, ECX sets daily reference price so that traders will use this information to make decision. Today's reference price is yesterday's last closing price and it is calculated for the whole coffee contracts. If a contract is not sold one day after a trade, the same problem occurs similar to the above case. In both cases the market is required to determine the daily selling price of all contracts which requires more resource.

A trading session is defined as specified period of trading during which a single contract or group of contracts will be traded [49]. ECX trading floor holds many trade sessions for transacting different contracts in a day. In the case of coffee trading a trade session is a specific time determined by the market where one origin and one processing type coffee contracts are traded. One origin coffee in traded in two sessions; first washed coffee contract of that origin coffee are traded after it is finalized unwashed will continue. The starting and ending of one session is announced ECX trade administrator. In a given day the market undertakes about 32 different trading sessions for export coffee contracts. In each trading session the market displays on its price board the price limit for all coffee contracts to be traded in that session which requires much time for administering the trade. To illustrate how trades are matched on ECX trading floor: assume this is session for Limmu origin coffee; the trade administrator announces the starting of washed Limmu origin coffee;

- The price board displays the price limit for all washed Limmu contracts (18 contracts);
- Buyers and sellers of washed Limmu contracts enter into the trading floor;
- They transact their commodity by offering a grade, price, quantity, and warehouse.

- If a buyer wants to buy ten contracts of grade 3 of Limmu A coffee, at a price of 1000 birr, he would yell/ shout/ Limmu A grade 3 at 1000, at Jimma warehouse by showing one hand with the palm facing outward.
- If the seller wants to sell the same, he will yell the same quote in the same way;
- After combination of hand-signals and vocal representation a trade is executed.
- ECX will undertake the price and commodity reconciliation in a minute.
- After this session has finalized unwashed Limmu will be continued.

Among the major role of ECX is market data dissemination which means the exchange shall provide timely and complete price information to the public through variety of media [14, 49]. In the above example the weighted average price of all contract traded in that session will be disseminated as of the daily selling price of washed Limmu coffee; which fairly represents its actual price.

Furthermore, ECX is strategically thinking to start futures market. Futures market is an auction market in which participants buy and sell Future contracts. Futures contract is an agreement between two parties to buy or sell a specified commodity of standardized quantity for a price agreed today with delivery and payment occurring at a specified future date. This is typically traded at regulated commodity exchanges. In futures trading having one futures contract is mandatory. Kolb[37] summarizes the benefits of futures markets as follows: (1) price discovery where the market observers can form estimates of what a price of a given commodity will be at a certain time in the future by using the information contained in futures prices today. (2) Price risk management. The prices of commodities are discovered and disseminated continuously on commodity futures exchanges. The price discovered on the futures market provides an idea about the price that is likely to prevail at a future point of time. Farmers can protect themselves against undesirable price movements and decide upon cropping pattern. The merchandisers avoid price risk. Processors keep control on raw material cost and decreasing inventory values. Moreover, future markets enable to increase the market liquidity.

Coffee is a global commodity that is traded worldwide on a daily basis. New York board trade (NYBOT) is a futures market where washed Arabica coffee contracts represented by Coffee "C" are traded. Similarly, Robusta coffee is traded in London International futures market [8, 27].

In NYBOT, Coffee “C” contract is the world benchmark for all Arabica washed coffee produced in several Central and South American, Asian and African countries. Based on coffee quality NYBOT has five coffee grades depicted in Table 1.3 [7, 27].

Table 1.3: NYBOT coffee grade

Class	Defect	Description
Class 1	0-5	Get premium price than Coffee “C”
Class 2	6-8	
Class 3	**9-23**	**Coffee “C” - is traded in the market**
Class 4	24 -86	They are sold with discounts from Coffee “C”
Class 5	>86	

Only Coffee “C” representing all Arabica washed coffee contracts is traded and this price is disseminated worldwide. NYBOT uses this standard coffee contract COFFEE “C” to calculate other coffee grade price. If it is judged of higher quality it will receive a premium price. On the other hand, if it is inferior quality, it will be sold with a discount calculated based on the standard one [7, 8].

In futures trading having one futures contract is mandatory for ECX to start futures trading the market need to define one or few standard contracts like that of NYBOT and other futures market that can represent all other coffee contracts that are traded in the market.

This research aims to study a price relationship between ECX coffee contracts and to propose a predictive model that can predict the daily selling price of other coffee contracts. As a result the market will trade only one contract and other contracts daily selling price will be determined based on it. To the best knowledge of the researchers so far no research was conducted on ECX coffee contract relationship. Moreover, ECX coffee grade classification method is unique to ECX.

1.1. Statement of the Problem

- Currently, Ethiopia Commodity Exchange defines about 32 trading sessions in a given trading day for different coffee origins and processing types. It utilizes much time and human effort for administering the trade. In addition, the traders have to wait until the session that they want is started.
- Setting price limit for each coffee contract in all trading sessions requires much resource.
- The market data disseminated by ECX fairly represents real commodity price because it is the weighted average of many coffee contracts that are traded in one trading session unlike that of coffee "C".
- ECX is strategically thinking to start futures trading. In futures market one or few internationally accepted contract that can represent and predict the price of other coffee contract must be modeled.

1.2. Significance of the Study

The result of this study will benefit ECX and its stakeholders in the following major aspects:-

- ECX will have efficient trading activity. The trading operation will be manageable and easily administered because ECX will trade only one contract the market is required to define one trading session. Moreover, setting reference price for one contract will be simple.
- ECX will disseminate actual coffee contract price.
- ECX the surveillance activity will be simplified because the market is only required to study and analyze the price behavior of one contract and the factors that are associated with this contract.
- The output of this study will be an input for ECX futures trading plan.

1.3. Objective of the Study

1.3.1. General Objective

The general objective of this research is to examine the price relationship between ECX coffee contracts and some other independent variables and to propose a predictive model that represents the required relationship.

1.3.2. Specific Objectives

To accomplish the general objective, the following specific objectives will be carried out.

- Conducting review of literature on data mining techniques.
- Identifying modeling techniques that are appropriate for this prediction task.
- Preprocessing the dataset.
- Conducting experiment to build a prediction model.
- Evaluating the performance of the models and select the best model for future prediction.

1.4. Scope and Limitation

ECX has 22 coffee origins; totally the market executes trade for about 421 coffee contracts. In the last five and half years from 1/1/2009 to 30/6/2014 ECX has transacted 409,530 lots[6] of coffee. From this amount of export coffees which are dedicated for the international market account for 83% while local coffees that are utilized for local consumption accounts only 11% of total. The study focuses only on export coffee contracts of all origins and all grades.

Moreover, export coffee contracts that are traded for very limited number of days in a given period of time are not considered in the study.

[6] Lot:- he standard lot size of an ECX contract is 50 quintals or 5 tons; it is designed to be suitable to current conditions of small truck transport in rural Ethiopia. One lot refers net weight of 60 kg for washed coffee or net weight of 85 kg for unwashed coffee

1.5. Research Methodology

1.5.1. Literature Review

Literature review has been conducted to assess concepts, techniques and applications of data mining technology and to get domain knowledge about the problem. In order to get a deeper understanding of how ECX is performing its trading activity ECX procedures, relevant documents and ECX website were reviewed.

In order to select modeling techniques that best suit the problem different data mining books, research works, journals and published articles on the application of data mining in commodity market and stock market prediction were reviewed. After reviewing different predictive model, the researchers select two artificial neural network methods: the multilayer perception feed forward neural and radial basis function for the study.

1.5.2. Data Mining Process Model

There are different data mining process models; the popular process models are SEMMA, CRISP-DM and KDD. In this study, we have used the CRISP-DM (Cross-Industry Standard Process for Data Mining). This methodology was proposed in the mid 1990s by an European consortium of companies to serve as a non-proprietary standard process model for data mining and it is widely applied process model. It consists of the following six steps: business understanding, collecting and understanding the data, data preparation, modeling, evaluating the model by using well known evaluation methods and finally deploying the model to predict future values [6].

Business Understanding

Understanding of the requirements from a business perspective is the core component in addressing the research objective. In this regard, the researcher has conducted discussion with the domain experts and analyzed relevant documents. This lets the researcher to get a deeper understanding of the business, to define the problem and articulate the goals of the study.

Data Collection and Data Understanding

Data understanding begins with an initial data collection and continues with tasks that make the researcher familiar with the data. Data understanding includes collecting the data, selecting the relevant attributes from the data and analyzing the data to address the problem.

The potential source of data for this study is ECX database. The data set that will be used for this research is ECX daily coffee trade data. From 1/1/2009 to 4/30/2013 daily coffee price data were collected and employed for analyzing the problem, for model building and testing the goodness of the models.

Data Preparation

The quality of data affects the data mining result and thus to improve the quality of data, the data need to be preprocessed. The data preparation (sometimes called as data preprocessing) is useful to make the collected data suitable for the selected data mining software. The data preparation process includes data cleaning, data summarization, data integration, data reduction, feature selection and finally preparing the data into a form that is acceptable to the neural network [3].

Model Building

There are a variety of machine learning techniques which are used to build a predictive model. To select an appropriate modeling technique, reviews were made on the application of machine learning in developing predictive model. Specifically on commodity price prediction, stock market prediction and function approximation problems.

Supervised machine learning techniques are used to build predictive models using the daily trading data; the models should learn the relationships that exist between coffee contracts. Finally, the researchers have selected artificial neural network to construct the model. The two well known and widely applicable methods for representing relationship and building predictive model the multilayer perception with back propagation algorithm and the radial basis function are used.

Model Evaluation

The model should be evaluated thoroughly to be certain that it properly achieves the business objectives; the coefficient of determination (R^2) and Mean Squared Error (MSE) are the most commonly used performance matrix in prediction problems. In this study the performance of each constructed models were evaluated by R^2 and MSE and the goodness of each model was compared and a model with best performance is selected for future prediction process.

1.5.3. Tools

The tools used in this study for preprocessing and analyzing the dataset are Microsoft Office Excel and STATA.

There are various tools that are used for building prediction models. Some of them are R, SPSS Modeler, Weka, SAS Predictive Modeling and Matlab. For this study, Matlab 2013a was used for building predictive models. The tool is selected because it is widely applied in prediction problems. Moreover, it is flexible and easy to apply.

Finally, Microsoft Office Word application was used for document preparation and Microsoft Office PowerPoint presentation was used to present the result of the study.

1.6. Organization of the Thesis

The thesis is organized into six Chapters, including the current one. Chapter Two presents the literature review and review of related work. The literature review briefly explains about data mining techniques, machine learning, artificial neural network and the application of machine learning in stock market and commodity market. Chapter Three discusses the data analysis of the study in detail. Chapter Four presents the design of the system. Chapter Five presents the experimental set up and the experimental result. The last Chapter presents the conclusions and recommendations of the study.

CHAPTER TWO

LITERATURE REVIEW AND RELATED WORK

A number of published papers; thesis works and web sites were reviewed on different modeling techniques. This chapter explains about data mining techniques, machine learning, artificial neural network and applications of ANN.

2.1. Overview of Data Mining

Data mining is defined as the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [22]. The field of data mining has been growing rapidly due to its broad applicability, achievements and scientific progress.

Data mining is a crucial step in the discovery of knowledge from large datasets. Mining process is more than data analysis which includes classification, clustering, prediction and association rule discovery.

Data mining involves an integration of techniques from multiple disciplines such as database and data warehouse technology, statistics, machine learning, high-performance computing, pattern recognition, neural networks, data visualization, information retrieval, image and signal processing, and spatial or temporal data analysis [22, 38].

These days a number of data mining applications have been successfully implemented in various domains such as customer relationship management, web applications, manufacturing, retail, finance, banking, monitoring, surveillance, teaching support, climate modeling, fraud detection, stock market, health care, telecommunication, and risk analysis.

2.1.1. The Data Mining Task

Data mining is applicable in predictive modeling, descriptive modeling and exploratory data analysis, discovering pattern and rules [48]. Each of these applications will be explained in brief as follows.

Predictive Modeling

It is building a model for the dependent variable from one or more independent variables. The value of the dependent variable will be predicted from the known values of other variables. Time series prediction in finance and other fields and function approximation are some of the applications areas. In predictive modeling one identifies patterns found in the data to predict future values. Classification and regression are two forms of data analysis that can be used to predict future data [22].

Classification methods create classes by examining already classified cases and inductively finding the pattern typical to each class. Regression uses the historical relationship between an independent and a dependent variable to predict the future values of the dependent variable. The difference between classification and regression is the type of output that is predicted; classification predicts class membership, whereas regression models continuous valued functions [22]. Many businesses use regression to predict future sales, stock prices and currency exchange rates.

Descriptive Modeling

It describe all the data, it includes models for overall probability distribution of the data and groups and models describing the relationships between the variables. Some of the commonly used descriptive modeling techniques are clustering and data visualization.

Clustering is used to identify a finite set of categories or clusters to describe the data. It involves partitioning data according to natural classes present in it, assigning data points that are "more similar" to the same "cluster". In clustering, no data are tagged before being fed to a function. The goal of clustering is to sift/filter the data to produce a segmentation of the input records. Different clustering functions will hence yield different sets of sorted data. It is up to the miner to determine what meaning, if any, to attach to the resulting clusters [3].

Visualization method is another powerful form of descriptive data mining [3]. It is a means for presenting data, both at the input and the output stages. Visualization techniques may help to discover relationships between features at the input stages, and explain the data mining results present them to the decision makers at the output stage.

Discovering Patterns and Rules

It is concerned with pattern detection. Some of its applications are handwritten recognition and face recognition [46].

2.1.2. Data Mining Process Models

To date many data mining and knowledge discovery process models have been developed. The most used in scientific research works, in industrial and academic projects are SEMMA, CRISP-DM and KDD.

SEMMA (Sample, Explore, Modify, Model, and Assess) was developed by the SAS Institute. The SEMMA process offers an easy to understand process, allowing an organized and adequate development and maintenance of DM projects. The SEMMA analysis cycle guides the analyst through the process of exploring the data using visual and statistical techniques, transforming data to uncover the most significant predictive variables, modeling the variables to predict outcomes, and assessing the model by testing it with new data [30].

KDD (Knowledge Discovery in Databases) is generally used to refer to the overall process of discovering useful knowledge from data, where data mining is a particular step in this process [22]. The additional steps in the KDD process, such as data preparation, data selection, data cleaning, and proper interpretation of the results of the data mining process, ensure that useful knowledge is derived from the data

CRISP-DM (Cross-Industry Standard Process for Data Mining) was developed in 1996 by analysts. It is applicable in typical data mining problems such as data description and summarization, segmentation, concept descriptions, classification, prediction, dependency analysis [6]. CRISP-DM is now being used in industry as the standard for a technology-neutral data mining process model [30].

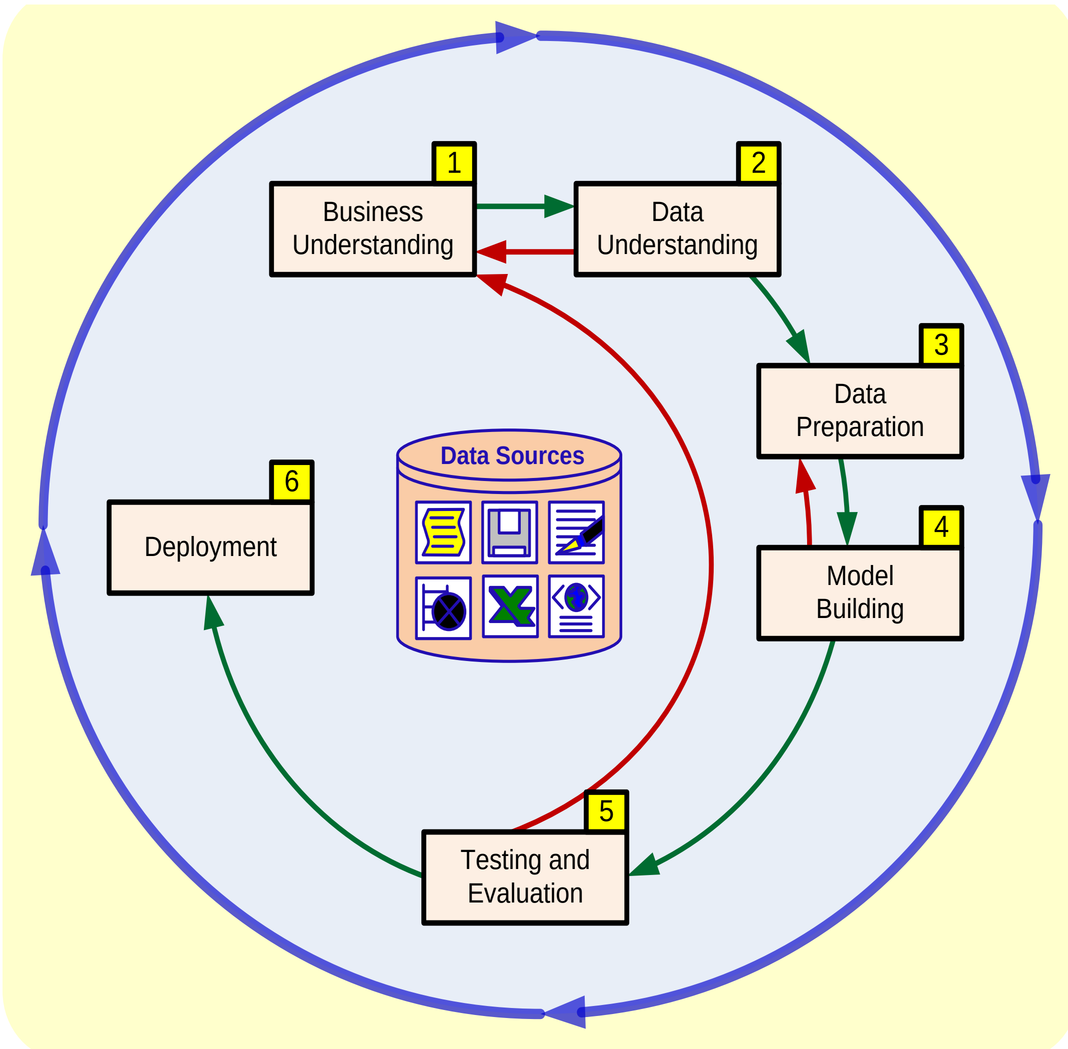


Figure 2.1: The CRISP-DM life cycle

As indicated by [6] the model consists of a cycle that comprises six stages: presented in Figure 2.1. All stages of CRISP data mining model are properly organized, structured and defined allowing that a project could be easily understood or revised [38].

According to the survey results of (KdNuggets.com, 2002; KdNuggets.com, 2004; KdNuggets.com, 2007a), CRISP-DM is the most widely used data model and the researchers selected this data model for this study.

2.2. Machine Learning

Machine learning, a branch of artificial intelligence, concerned with the design and development of algorithms and techniques that allow computers to learn. The major focus of machine learning research is to extract information from data automatically, by computational and mathematical methods.

According to Ticlavilca *et al.* [51] machine learning (ML) theory is related to statistical inference wherein a model is capable of learning to improve its performance on the basis of its own prior experience. Machine Learning includes a number of advanced statistical methods for handling regression and classification tasks. ML models include the artificial neural networks, Support Vector Machines and relevance vector machines [26, 51].

Currently, machine learning is applicable in many fields of studies. Some of the application areas are handwritten recognition, optical character recognition, image retrieval, medical diagnosis, bioinformatics, stock market prediction, robotic, analyzing social patterns, financial economics modeling and forecasting agricultural commodity prices.

Ticlavilca *et al.* [51] in their research work declare that due to the increased volatility in agricultural commodity prices increases the difficulty of forecasting accurately their price by using the simple or statistical methods less reliable and accurate in this new market environment. To overcome these limitations, machine learning models can be used as an alternative way to complex forecast models.

2.3. Artificial Neural Networks

Artificial neural networks (ANN) have been developed as generalizations of mathematical models of biological nervous systems. The first wave of interest in neural networks (also known as connectionist models or parallel distributed processing) emerged after the introduction of simplified neurons by [41].

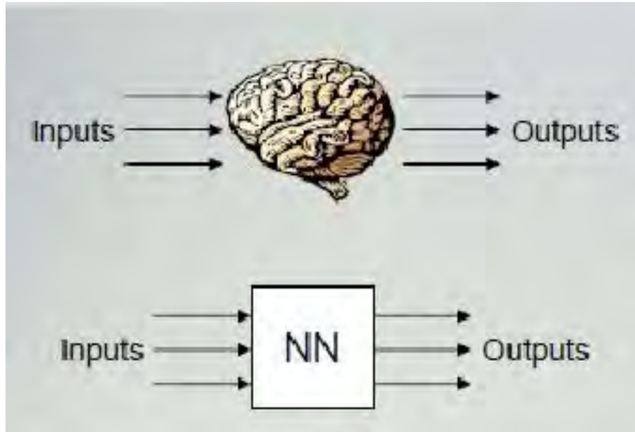


Figure 2.2: Artificial Neural Network and Brain Analogy

Figure 2.2 shows the similarity between artificial neural networks and the brain. The basic processing elements of neural networks are called artificial neurons, or simply neurons or nodes. A neuron is a special biological cell that processes information from one neuron to another neuron with the help of some electrical and chemical change. It is composed of a cell body and two types of out reaching tree like branches: the axon and the dendrites [31].

Mostly it has three layers. The first layer in the artificial neural network is the input layer which helps to accept external data into the artificial neural network. The second layer is the hidden layer where the input neurons are processed to convert to the output neurons. The last layer is the output layer relays the information out. The other important component in artificial neural networks is connections. These connections facilitate data transformation to each layer.

2.3.1. Advantage of using Artificial Neural Network

Artificial Neural Networks are particularly well suited to finding accurate solutions in an environment characterized by complex, noisy, or partial information. Several distinguishing features of ANNs make them valuable and attractive in prediction problems. The first important feature of ANNs is that they are considered to be universal functional approximator; they are able to approximate any continuous function to any desired accuracy [10, 18, 24, 25 29, 56] The second is its generalization ability; producing outputs from unseen inputs through captured patterns in previously learned inputs, what makes these techniques excellent classifiers and regression models [35]. This feature is extremely useful in financial forecasting, since the underlying relationships of the financial market are often unknown or hard to describe [55].

2.3.2. Taxonomy of Neural Network Architecture

ANNs are broadly classified as feed forward and feedback networks (FFNN) depicted in

Figure **2.3**. Feed forward networks where the data flow from input to output units is strictly feed forward manner. The data processing can extend over multiple (layers of) units, but no feedback connections are present, that is, connections extending from outputs of units to inputs of units in the same layer or previous layers [48].

Recurrent neural network (RNN) is a type of neural network which contains feedback connections with back-loop in it. Recurrent neural network has advantage over FFNN in that it saves patterns with time.

The well known learning algorithms, their architecture and their major application areas are presented in Table 2.1 [31].

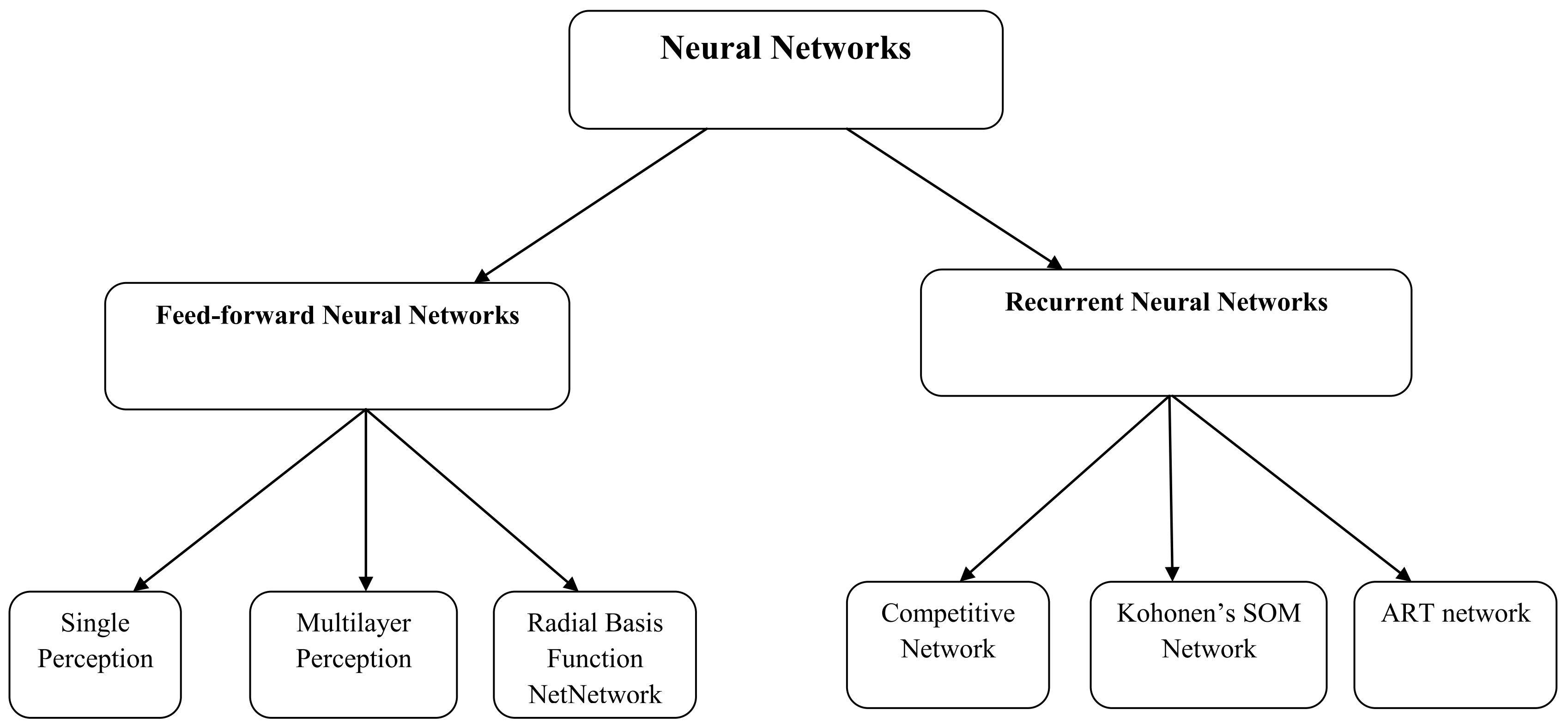


Figure 2.3: Classifications of ANN

Table 2.1: Well known learning algorithms, their architecture and major application areas

Paradigm	Learning rule	Architecture	Learning algorithm	Task
Supervised	Error correction	Single or Multilayer perception	Perception learning	Pattern classification
			Back propagation	***Function approximation***
			Adaline and madalin	Prediction and control
	Boltzmann	Recurrent	Boltzmann learning	Pattern classification
	Hebbian	Multilayer feed-forward	Linear discriminate analysis	Data analysis
				Pattern classification
	Competitive	Competitive	Learning vector quantization	Within class categorization
				Data compression
		ART network	ARTMap	Pattern classification
				Within class categorization
Unsupervised	Error correction	Multilayer feed-forward	Sammon's projection	Data analysis
	Hebbian	Feed-forward or competitive	Principal component analysis	Data analysis
				Data compression
		Hopfield Network	Associative memory	Associative memory
	Competitive	Competitive	Vector quantization	Categorization
				Data compression
		Kohonen's SOM	Kohonen's SOM	Categorization
				Data analysis
Hybrid	Error correction and competitive	RBF network	***RBF learning algorithm***	Pattern classification
				Function approximation
				Prediction and control

2.3.3. Applications of Neural Networks

Jain [31] and Kecman [35] summarize the major application areas of ANN as:-

- Pattern recognition: some of well known applications include character recognition and speech recognition.
- Clustering and categorization: data compression and data explanatory analysis.
- Function Approximation: various engineering and scientific modeling problems.
- Prediction: stock market prediction, weather forecasting are some of the applications.
- Time series forecasting, medical diagnostics, robotics, industrial process control, optimization, and signal processing.

2.4. Function Approximation by Neural Networks

Function Approximation is one of the most general uses of artificial neural networks. The general framework of the approximation problem is as follows. One supposes the existence of a relation between several input and one output variables. This relation being unknown, one tries to build an approximator between these inputs and the output. The structure of this approximator must be chosen and the approximator must be adjusted as to best represent the input-output dependency [39].

There are various data mining tools that have been used in approximating linear and non-linear functions. Some of them are ordinary least squares, non-linear least squares regression, nonparametric regression and ANN [33].

ANN can approximate any linear and non-linear function to any desired degree of accuracy; it can learn any linear, non-linear time series patterns and can extrapolate linear and non-linear patterns. This process involves two phases known as the learning phase (training) and the testing phase (forecasting). Neural networks have been mathematically shown to be universal approximator of functions [18, 23, 39].

The novelty of ANN lies on its ability to discover non-linear relationship without prior knowledge about relationship between the input and the output [5]. There are many different artificial neural network models currently applicable for prediction problems. The MLP and RBF are the most popular and influential models [55]. Recently, feed forward neural networks such as multilayer perceptrons developed by Werbos and Rumelhart and the radial basis function networks developed by Verleysen are used in approximating non-linear model and gives very good results [39].

2.5. Multilayer Perceptron Neural Network Model

Multilayer perceptron neural network is arranged in a layer of neuron. MLP network has three layers the input layer, the hidden layer and the output layer (Figure 2.4).

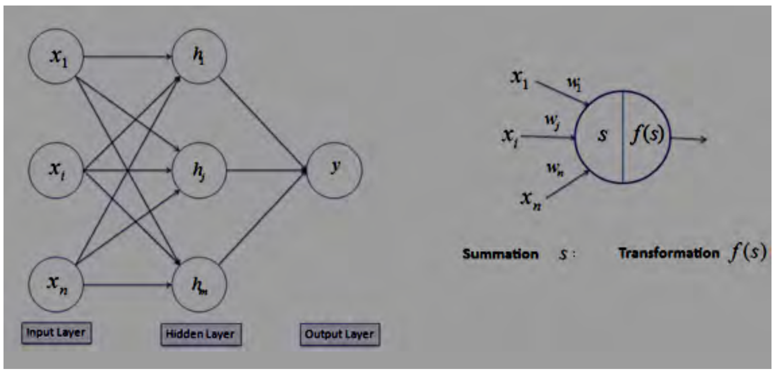


Figure 2.4: The basic MLP neural network architecture

The input layer has Xn input neurons, it distributes input signal to the hidden layer. The hidden layer has hm neurons and the activation function. Each neuron in the hidden layer sums up the input neurons with their corresponding weights and this result is passed to the transfer function (f(s)). The transfer function modifies the input to give a desired output. The output layer has one output neuron that is y. Each layer is connected in a feed forward manner. Choice of the number of the hidden layers, hidden nodes and type of activation function plays an important role in model constructions process.

2.6. Radial Basis Function Neural Network

A radial Basis Function neural network (RBF) is a special class of feed-forward network [42, 52]. Park and Sandberg [45] proved the universal approximation ability of RBF networks. The radial basis function neural network with sufficient number of hidden in the hidden layer is capable of approximating any given non-linear function to a desire degree of accuracy.

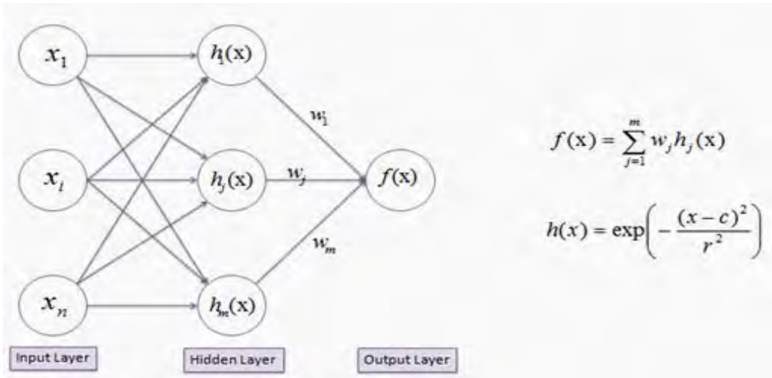


Figure 2.5: The basic RBF neural network architecture

The RBF neural network comprises three layers; the input layer, hidden layer and output layer. Between the input and output layers is a layer of processing units known as hidden units. Each of these implements a radial basis function neural network [4, 52]. Figure 2.5 depicts the basic architecture of RBF network. It has Xn input nodes in the input layer, Wn are the weights, hm(x) is the Gaussian activation function with the parameters r (the radius or standard deviation) and c (the center or average taken from the input space) defined separately at each RBF unit. The learning process is based on adjusting the parameters of the network to reproduce a set of input-output patterns. There are three parameters; the weight w between the hidden nodes and the output nodes, the center c of each neutron of the hidden layer and the unit width r. And f(x) is the output of the network.

Determination of the number of neurons in the hidden layer is very important because it affects the network complexity and the generalizing capability of the network. To build RBF network two parameters, the center and spread should be selected appropriately. The position of the centers in the hidden layer also affects the network performance considerably, so determination of the optimal locations of the centers is an important task. The spread value is selected based on the minimum error criteria of the developed model. Spread value determines the diversity of radial basis function.

2.7. Related Work

Traditionally, statistical method Autoregressive integrated moving average (ARIMA) models have been used to model linear time series data but they cannot capture the non-linear pattern. In the last two decades artificial neural networks have been proposed as an alternative to the traditional forecasting method in particular for the presence of non-linear data pattern [50].

There has been growing interest in using of artificial neural networks to analyze the historical data and make predictions in many domains. ANN applications in finance were reviewed by Fadlalla and Lin [16] and they are applied in commodity market, bankruptcy prediction, stock market forecasting and bond rating. The related works were reviewed; in particular to commodity and stock market prediction.

Artificial neural networks are also used for predicting agricultural commodity prices. Kohzadi *et al.* [36] uses multi layer perception neural networks with back propagation and ARIMA to model monthly live cattle and wheat prices and compare the results of the two models. The multi layer perception neural networks are better in capturing the pattern and improving the prediction performance. Similarly, a neural network was modeled to recognize and direct a buy/sell pattern for the live cattle commodities futures market [9]. Shahwan and Odening [50] used a hybrid of ANNs and ARIMA model to predict agricultural commodity prices and the hybrid model performs higher accuracy result than ARIMA model. Zou *et al.* [56] have used feed forward neural network (FFNN) and time series models for Chinese food grain price forecasting. FFNN model was found out to be the best.

Moreover, ANNs have been used in stock market prediction by many researchers. Naeini *et al.* [43] suggested a predictive model based on MLP neural network for predicting stock market changes in Tehran Stock Exchange Corporation. The model uses a three layer neural network in which the input layer has 3 neurons, the hidden layer there are h neurons which are fully connected to the input and output layers. There is one neuron in output layer which predicts the expected stock value of the next day of the stock market. Using this model, they can predict the next day stock value of a company only based on its stock trade history and without any information of the current market with a prediction error of around 1.5%.

Yildiz *et al*. [54] developed an efficient three layer neural network with back propagation algorithm to predict the direction of Istanbul Stock Exchange National-100 Indices and the model forecast and direction of ISE National -100 to an accuracy of 74.51%.

Kara *et al*. [32] attempt to develop two efficient models; ANN and support vector machine (SVM) and compared their performances in predicting the direction of movement in the daily Istanbul Stock Exchange (ISE) National 100 Index. Their experimental results showed that average performance of ANN model (75.74%) was found significantly better than that of SVM model (71.52%). They use the computational data mining methodology to predict seven major stock market indexes in Tehran stock market. Two learning algorithms linear regression and feed-forward neural network with back propagation used to determine and explore the relationship between some variables as independent factors and the level of stock price index as a dependent element in the stock market under study over time. The performances of the prediction models were evaluated and feed-forward back propagation algorithm resulted in better prediction accuracy. The related works and the method applied to build prediction model are presented in Table 2.2.

Table 2.2: Summary of Related works

I.No	Authors	Title	Objective and Result	Method
1	Aghababaeyan *et al*. [2]	Forecasting the Tehran Stock Market by Artificial Neural Network	✓ To predict the major stock prices indexes in Tehran Stock Exchange the authors' have used two learning algorithms Linear Regression and feed-forward back propagation ✓ FFNN algorithm resulted in better prediction accuracy.	• Linear Regression • Feed-forward back propagation
2	Yildiz *et al*. [54]	Forecasting the Istanbul Stock Exchange National 100 Index Using an ANN	✓ The purpose of the study was to predict the direction of the Istanbul Stock Exchange National 100 Indices using ANN model. ✓ The result of this study was the model forecast the direction of the ISE National-100 to an accuracy of 74.5%.	• Three-layer neural network with Back propagation algorithms

I.No	Authors	Title	Objective and Result	Method
3	Kara *et al*. [32]	Predicting direction of stock price index movement using artificial neural networks and support vector machines Istanbul Stock Exchange	✓ To develop and compared the performances of ANN and SVM in predicting the direction of movement in the daily Istanbul Stock Exchange (ISE) National 100 Index. ✓ Their result was ANN model perform better than that of SVM model.	• ANN • SVM
4	Perwej and *Perwej* [47]	Prediction of the Bombay Stock Exchange (BSE) Market Returns Using ANN	✓ In this paper, they investigate to predict the daily excess returns of Bombay Stock Exchange indices over the respective Treasury bill rate returns. They have applied different methods to predict the excess return time series using lagged value. ✓ ANN model is superior compared to the autoregressive models because they are able to capture not only linear but also non-linear patterns in the underlying data.	• Autoregressive • Feed forward Artificial Neural Networks
5	Zou *et al*. [56]	An investigation and comparison of ANN and time series models for Chinese food grain price forecasting.	✓ They build time series models for Chinese food grain price forecasting. ✓ FFNN model was found out to be the best	FFNN
6	Neeraj *et al*. [44]	Artificial neural network models for forecasting stock price index in Bombay stock exchange	✓ Their objective was to model of the Indian stock market data using NN. They studied the efficiency of ANN in modeling the Bombay Stock Exchange	ANN
7	Budhani *et al*.[5]	Application of neural network in analysis of stock market prediction	✓ Neelima used FFNN with back propagation algorithm for training the data to make prediction.	FFNN

I.No	Authors	Title	Objective and Result	Method
8	Ebrahimi *et al*. [12]	Prediction of agricultural commodity price using ANN: case of chicken price in Fars province, Iran;	✓ Price of chicken was predicted using different methods; ✓ They found that FFNN are more efficient than the other methods in predicting the chicken price.	Autoregressive conditional hetroskedasicity, Autoregressive moving average and FFNN.
9	Adebiyi *et al*. [1]	Stock Price Prediction using Neural Network with Hybridized Market Indicators	✓ Their objective was to predict the daily stock price of the Nigerian stock market by FFNN with back ward propagation. ✓ They proved that the Nigerian stock market can be predicted by FFNN.	FFNN

2.8. Summary

Machine learning is the design and development of algorithms and techniques that allow computers to learn. The major focus of machine learning research is to extract information from data automatically, by computational and mathematical methods. The commonly used machine learning model is artificial neural network.

Artificial Neural Networks are inspired by the human brain, they are able to learn and generalize knowledge from previous events to predict future unseen events [35]. Artificial neural networks are applied in a larger domain of problems because of their nonlinear system modeling capacity. ANNs are able to map the relationship between input and output; they can learn the relationship stores it in its parameters and it uses this knowledge to predict new data.

The major advantages of ANN over the traditional method are:-

- ANNs have good generalization ability. Once the network in trained by training data set it can correctly infer the unseen data even if the data contains noisy information.
- ANNs are universal function approximator. ANNs can extract relationship between the input and output of the system even if this relationship is very complicated. It can model data with nonlinear non-linear pattern flexibly.

- Its prediction accuracy is generally high and robust, works even the training data have errors.

The MLP is selected because it is the most commonly used in time series and financial forecasting [33]. Moreover, it is proved that an MLP neural network can approximate any complex continuous function that enables us to learn any complicated relationship between the input and the output of the system and it has a simple architecture and relatively easy to implement [33]. While the RBF networks model is selected because it have good approximation capabilities, simpler network structures and faster learning algorithms. Due to this characterstics RBF networks have been widely used in many science and engineering applications [45]. Some of the application areas are function approximation, curve fitting, time series prediction, control and classification.

Agricultural commodity price data are known for their non-linear pattern. Traditional prediction models are unable to model these data sets. From the review of related works, ANNs offer a promising alternative approach to traditional linear methods because they are capable of performing non-linear modeling without prior knowledge about the relationship between input and output variables. Among ANNs the multi layer perceptions neural network trained by back propagation algorithm and the radial basis function neural network were most applied techniques for prediction problems and both are known for their universal function approximation ability. From the review of literatures and related works the researchers have decided to use the multilayer perceptron neural network and the radial basis function neural network methods to predict the daily selling price of ECX coffee contracts.

CHAPTER THREE

DATA ANALYSIS

This chapter presents the three stages of CRISP data mining model business understanding, data understanding and data preparation.

Business understanding focuses on understanding the business problem from the business perspective and then converting the knowledge into a data mining problem definition. The business understanding task includes determining business objectives; determining data mining goals and produce plan to solve the existing problem.

While data understanding starts with an initial data collection and proceeds with activities in order to get familiar with the data, describing the data, exploring and verifying the quality of data [8].

Data in its raw form is not always good for analysis, as real-world data tend to be incomplete, noisy, and inconsistent. One of the most important tasks in data mining is preparing the data in a way that is suitable for the specific data mining tool and thus the data will be appropriate for model building process [22].

3.1. Business and Data Understanding

The success of designing neural network depends on clear understanding of the problem [33]. The first objective of the researcher in the data mining stream is to understand the business perspective thoroughly. A possible consequence of neglecting this step is to expend a great deal of effort producing the right answers to the wrong questions [30].

In this study, in order to get a deeper understanding of existing ECX trading system (the trading activity and price quotation) and coffee trading data relevant documents and ECX website were reviewed and discussion were made with ECX domain experts in addition the researcher has observed ECX coffee trading activity.

3.1.1. Data Collection

Mostly, the frequency of the data depends on the objective of the researcher. Many studies on commodity and stock markets use daily data to design neural network [33]. The data used for this research is ECX daily coffee trading data. The data source is ECX market data dissemination department. For this study, five years and 6 months data from 1/1/2009 to 30/6/2014 daily coffee trade data were used for model building and testing the goodness of the model.

3.1.2. Descriptive Data Visualisation

To understand the general behavior of coffee trading data same relevant statistics are computed. The descriptive statistics of coffee contracts are presented in Annex 3.

In the past five year's period, a total of 409,530 Lot of different origin coffee have been traded in the market Table 3.1. From this amount of specialty and export coffees which are dedicated for the international market in aggregate account for 88% while local coffees that are utilized for local consumption accounts only 11% of total.

Table 3.1: ECX Coffee trade volume by market type in lot

Market type	Year						Grand Total	% share
	2009	2010	2011	2012	2013	2014		
Export	518	45,794	63,033	92,609	77,739	61,792	341,486	83%
Local	95	4,816	11,803	12,255	12,222	5,843	47,034	11%
Specialty	361	3,627	1,774	4,979	5,660	4,609	21,009	5%
Grand Total	973	54,237	76,610	109,844	95,621	72,244	409,530	100%

The average daily closing prices of the export coffee contracts are presented in Figure 3.1. The price does not follow regular pattern.

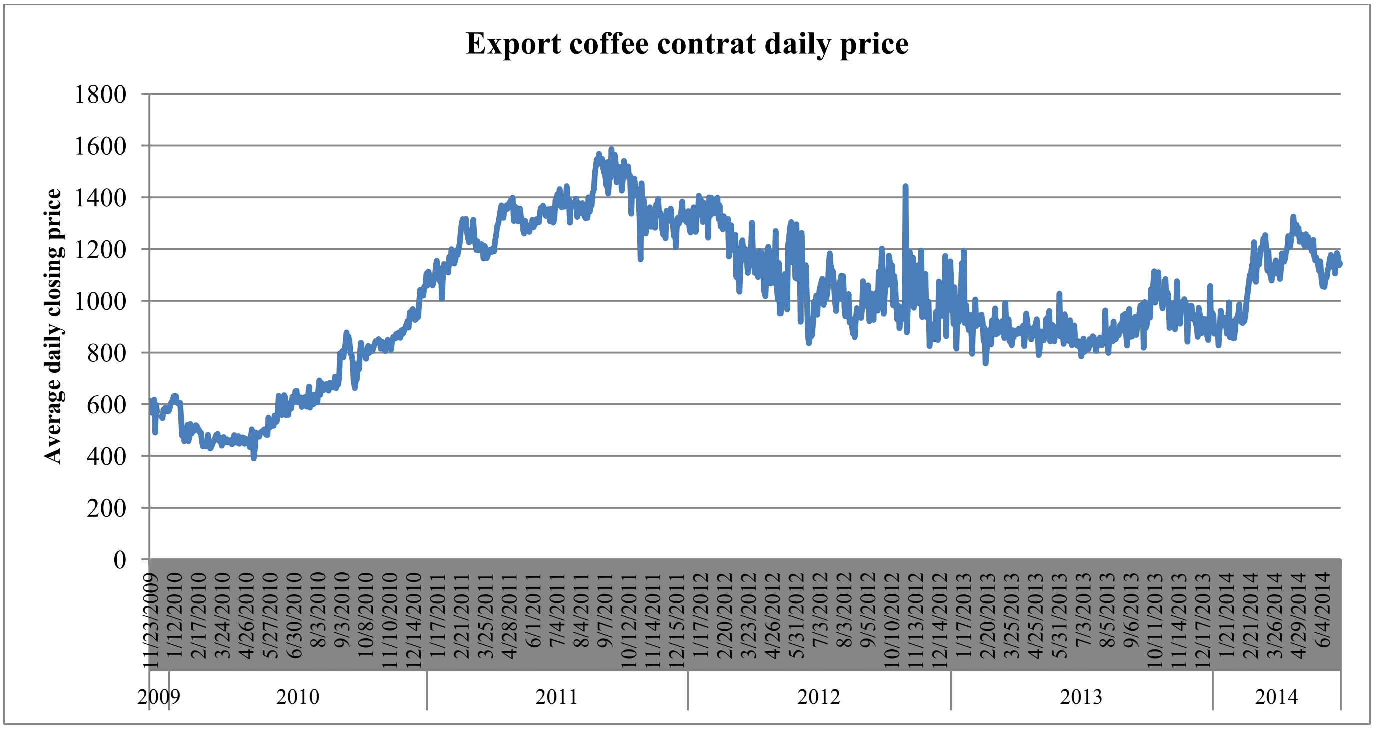


Figure 3.1: Export daily coffee Average closing price by market type

3.1.3. Data Preparation

Data preparation refers to analyzing and transforming the input and output to minimize noise, highlight important relationships and detect trends to flatten the distribution of the variable to assist the neural network in learning the relevant pattern [33].

The data preparation phase covers all the activities to construct the final dataset from the initial raw data. Data preparation tasks include attribute selection, data transformation, data cleaning and data construction [30].

To prepare the data, the first step is selecting the relevant attributes for the study. After the attributes are selected data cleaning process was continued. Finally, data transformation was employed to make the cleaned data more suitable for the selected neural network tool.

STATA and Microsoft Excel were used to compute important descriptive statistics in order to understand the general behavior of the dataset, to test whether there is a relationship between different coffee contracts and to find out the degree of relationship.

3.1.4. Attribute Selection

Selecting the relevant attributes will enable to handle the common problem that could exist during the prediction process, i.e. memory utilization and inefficiency in prediction. Relevance between attributes is usually measured by the degree of relationship between the target and other attributes. Removing the irrelevant attributes will improve the accuracy of the prediction model [28]. Therefore, the irrelevant attributes that do not contribute to the prediction task should be identified and removed. The trading data contains 23 attributes the description of total list of attributes is shown in Annex 2. Finally 9 attributes were selected to build the prediction model.

3.1.5. Description of the Selected Attributes

After the relevant attributes were selected from the database new dataset was created and stored. The newly created dataset incorporates about 10 attribute including the newly created target attributes and 44,485 records. The description of the selected attributes is presented in Table 3.2.

Table 3.2: Selected attributes and their and descriptions

I.No	Field Name	Description	Data Type	Possible Values
1	Tradedate	The date the trade is performed, coffee trade is performed from Monday to Friday	Date	Date/month/year
2	Origin	The origin of coffee	String	Sidama, Yirgachefe, Jimma, Lekemte, …
3	ProcessingType	Washed or unwashed	String	Washed, unwashed
4	MarketType	Export, local or specialty	String	Export
5	Class	The class where the commodity. …	String	WSDA, WSDB, WSDC
6	Contract	The coffee contracts washed Sidama A grade 3, contracts washed Sidama A grade 4, and so on.	String	WSDA3, WSDA4, WSDA5 and so on
7	ClosingPrice	The average daily price of a contract	Number	Price value normalized in to the range between -1 to 1.
8	Warehouse	The place where the coffee is stored	String	Warehouse information value normalized in to the range between -1 to 1.
9	Volume	Volume of coffee traded in the market; measured in lot	Number	2;10.5; 50, and more

3.1.6. Data Cleaning

Data data cleaning is a process that attempts to fill in missing values, smooth out noise while identifying outliers, and correct inconsistencies in the data [22]. Witten and Frank [53] described data cleaning as a time-consuming and labor-intensive procedure but it is absolutely necessary for successful data mining. Some of the data cleaning tasks that are applied in this study are removing outliers and handling missing value.

Handling Outliers

Outliers are extreme values that lie near the limits of the data range (extremely large or small values) or go against the trend of the remaining data. Identifying outliers is important because they may represent errors in data entry. Also, even if an outlier is a valid data point and not an error, certain statistical methods are sensitive to the presence of outliers and may deliver unstable results [38]. There are varieties of methods to handle the outliers that exist in the data; they are eliminating outliers, replace values as missing values, rescaling, and normalization of variables.

In this study, the only cause for the existence of outlier is encoding error. To identify these outliers the researcher used data filtering followed by visual inspection methods. The data that are very far large was identified and the records are ignored from the analysis.

Handling Missing Values

Han and Kamber [22] discussed in their book, some of the solutions used to avoid missing value from the dataset are ignore the tuple, filling in the missing value manually; replacing the missing attribute values by the same constant value, using the attribute mean for all samples belonging to the same class as the given tuple.

The main reason for the existence of missing values in this study is the contract might not be traded in a given day; when this scenario occurs all the information associated with this empty trade data was discarded from the analysis.

3.1.7. Selecting Best Contract

To select the best contract from export coffee contract categories, two parameters from the trading data have been used; the volume of coffee traded in the market and the availability of the contract in a year time. Finally, a contract that is transacted in a larger volume and available for a long period of time in a year will be selected to be used as a reference contract.

Volume of Trade

The volume of coffee traded in the market was taken as one parameter to select the best contract. Coffee classes that are traded more than 5,000 lot in the five years period are extracted and plotted.

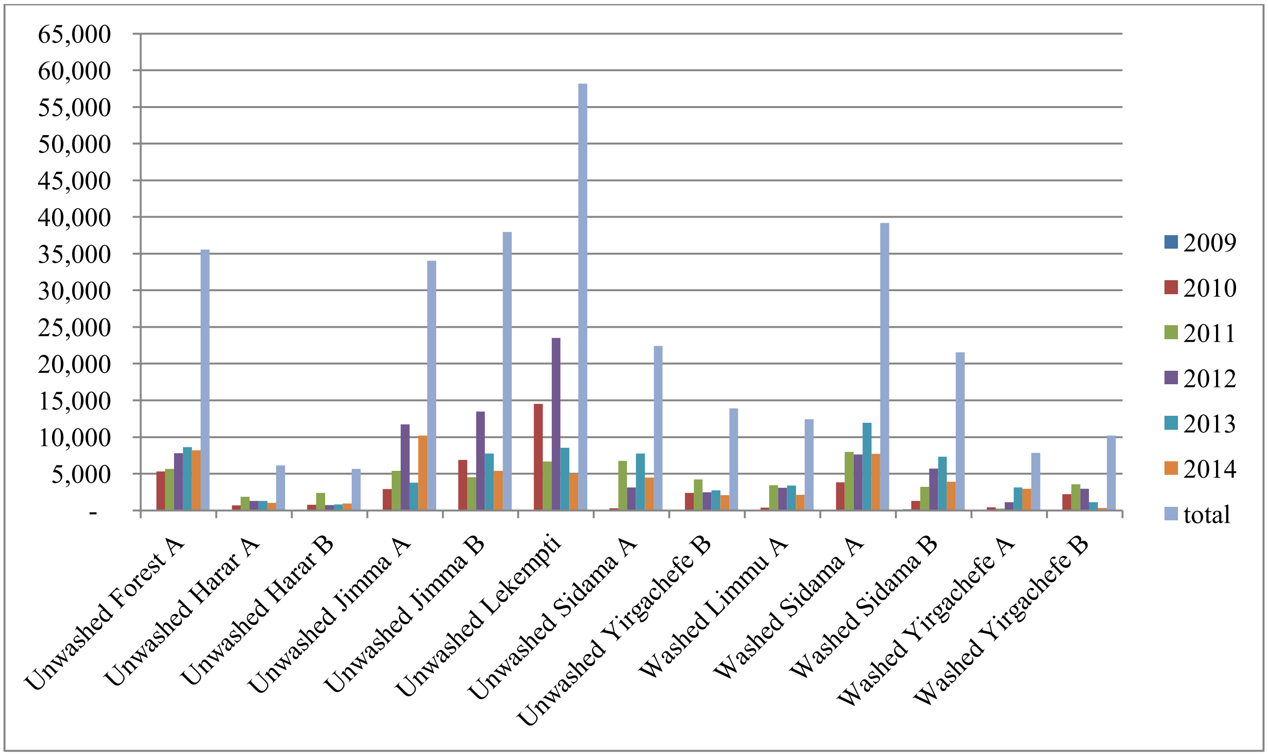


Figure 3.2: Volume of coffee traded in lot in coffee class

From Figure 3.2 coffee classes unwashed Lekempti, washed Sidama A, unwashed Jimma B and A, unwashed Forest A classes were traded in higher volume than others.

These coffee classes were selected to undertake detail contract level analysis which enables to identify the contract that is traded in a larger volume.

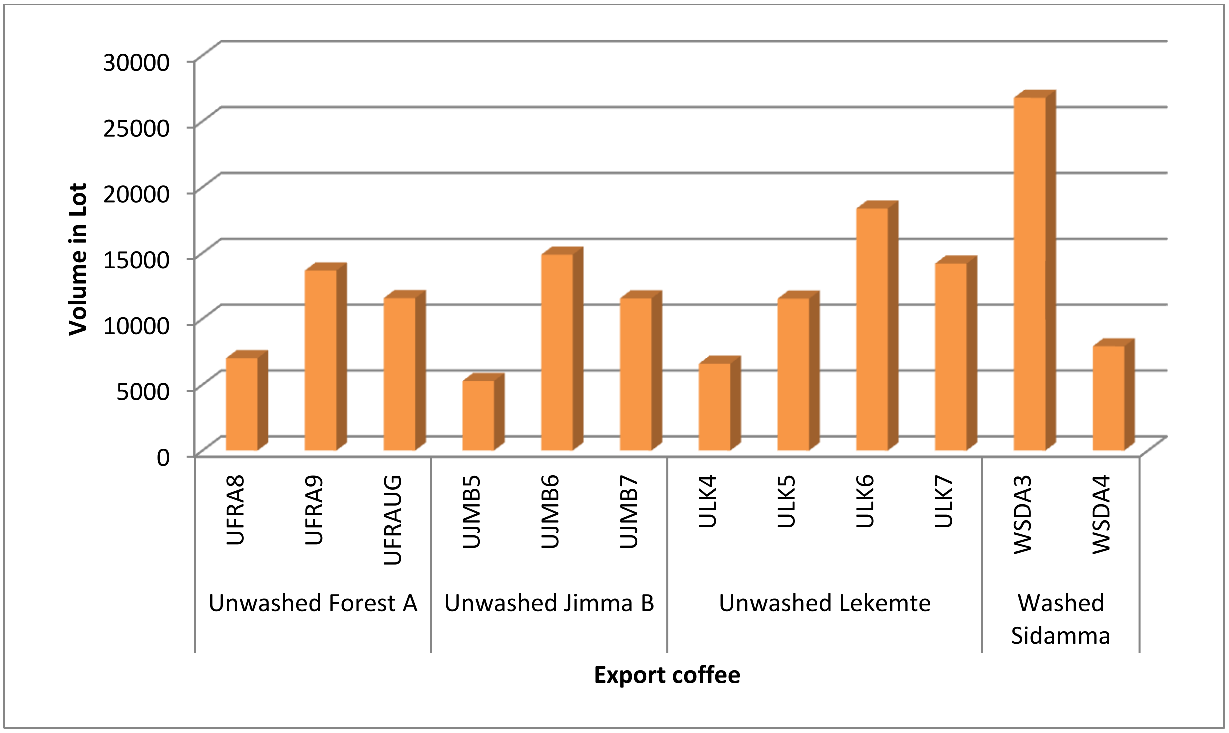


Figure 3.3: Coffee contracts traded more than 5,000 lots

From Figure 3.3 washed Sidama A grade 3 (WSDA3), unwashed Lekemte grade 6 (ULK6), unwashed Forest A grade 9 (UFRA9) and unwashed Jimma B grade 6 (UJMB) contracts were traded in a larger volume. Washed Sidama A grade 3 is the first in volume of trade it accounts about 8% of all coffee contracts traded in the last five years.

The availability of the contract

A contract is said to be available if it is traded in each trading day of the year. From the data, we have seen that from unwashed Jimma coffee contracts UJMA7, UJMA8, UJMB6 and UJMB7; from unwashed Lekemte contracts ULK6 and ULK7, from washed Sidama contracts WSDA3 and WSDB3 and from unwashed forest contracts UFRA9 and UFRAUG are traded in almost every trading day.

Based on the result of the above data analysis, WSDA3 was found to be the best contract in its volume of trade and availability and ULK6 is the second best contract.

3.1.8. Relationship Analysis

Here our interest is to examine price relationship between the selected coffee contract and other coffee contracts. The hypothesis is restructured as “The relationship between WSDA3 and other export coffee contracts can be modeled”.

To prove this hypothesis first the relationship between these coffee contracts should be investigated. Three scenarios are considered:-

Scenario I:- Relationship between WSDA3 and washed Sidama coffee contracts.

A reference contract WSDA3 is Sidama washed grade 3. To examine the price relationship within coffee contract in the same origin and same processing type to the reference contract i.e. relationship between WSDA3 and different grades of washed Sidama coffee contracts.

Scenario II:- Relationship between WSDA3 and unwashed Sidama coffee contracts

Relationship within same contract origins and different processing to the reference contract.

Scenario III:- Relationship between WSDA3 and other coffee origins

Relationship between WSDA3 and other coffee origins like Lekemte, Jimma, Yirgachefe, Limmu and of different processing type.

To discover the existence of relationship and type of relationship between the reference contract and other coffee contracts, the scatter plot matrix and the scatter plots were used. Moreover, to measure the degree of relationship between the contracts correlation coefficient matrix was computed by using STATA.

3.1.9. Scatter plot and Scatter Matrix

A scatter plot is one of the most effective statistical graph used to determine if there appears a relationship or pattern between two numerical attributes. The scatter plot is useful method for providing a first look at bi-variate data to explore the possibility of relationships [22].

Scatter plots of two data points (x, y) give us a summary:

- The direction of the pattern: is the relationship positive (x goes up and y goes up, x goes down and y goes down), negative (x goes up, y goes down), or is there no relationship?
- Is the relationship linear, quadratic, curved or something else?

A scatter plot matrix also helps us to explore relationship more than two numerical variables. It displays the relationship between two consecutive pair variables and the result is presented in a matrix.

Scenario I: - Relationship between WSDA3 and washed Sidama coffee

In Figure 3.4 the first row shows the response of WSDA3 against each washed Sidama coffee contracts. The scatter plot matrix depicts some linear and positive relationship between WSDA3 and other washed Sidama coffee contracts.

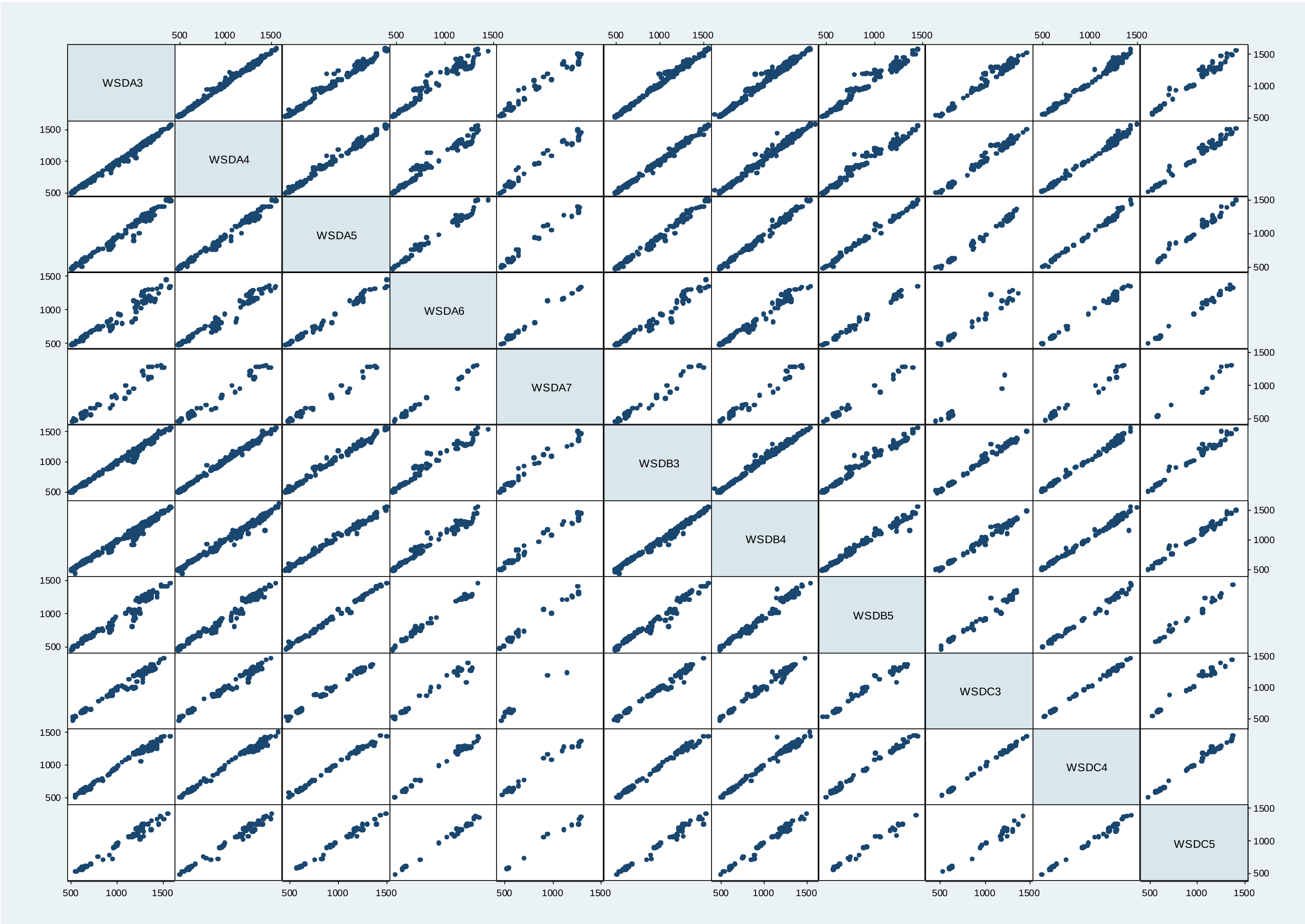


Figure 3.4: Scatter plot matrix between WSDA3 and washed Sidama coffee contracts

3.1.10. Correlation Coefficient Analysis

Correlation analysis measures the relationship between two quantitative variables. Correlation analysis helps to understand the independent variable's predictive abilities [30]. It is used to determine the existence and strength of relationship between two variables. A correlation is measured by a coefficient of correlation (r). It ranges from +1.0 to -1.0.

Correlation coefficient is given by:

$$r = \frac{\sum(Pi - \overline{Pi})(Pi' - \overline{Pi'})}{\sqrt{[\sum(Pi - Pi)^2][\sum(Pi' - \overline{Pi'})^2]}}$$

Where

Pi − is the daily selling price of the reference contract at specific trade date

$P\ i$ − is the daily selling price of the reference contract at specific trade date

If Pi and Pi are perfectly positively correlated, r = 1. If there is absolutely no relationship, r = 0 If they are negatively correlated, r = -1. As a rule of thumb, a strong correlation or relationship has **r**-value range from 0.85 to 1, or -0.85 to -1. In a moderate correlation, **r**-value ranges from 0.75 to 0.85 or, -0.75 to -0.85. In a weak correlation, not a very helpful predictor, **r** ranges from 0.60 to 0.74 or -0.60 to 0.74. Any relationship that has a correlation **r**-value that is 0.59 and below is not considered to be a reliable predictor.

The result of Table 3.3 shows that all washed Sidama coffee contracts are highly correlated with WSDA3. From the results of the scatter plot and correlation coefficient matrix, we have proved the existence of strong and positive relationship between the WSDA3 and all washed Sidama coffee contracts. And hence the first scenario of the research question is addressed. WSDA3 can represent all washed Sidama coffee contracts.

Table 3.3: Correlation coefficient matrix WSDA3 with other washed Sidama coffee contracts

CORR.	WSDA3	WSDA4	WSDA5	WSDA6	WSDA7	WSDA8	WSDAUGN P	WSDAUG	WSDB3	WSDB4	WSDB5	WSDB6	WSDB7	WSDB8	WSDB9	WSDBUGN P	WSDBUG	WSDC3	WSDC4	WSDC5
WSDA3	1.000																			
WSDA4	0.996	1.000																		
WSDA5	0.984	0.986	1.000																	
WSDA6	0.974	0.980	0.987	1.000																
WSDA7	0.979	0.982	0.976	0.977	1.000															
WSDA8	0.822	0.806	0.800	0.982	0.985	1.000														
WSDAUGNP	0.919	0.927	0.900	0.924	0.864	0.993	1.000													
WSDAUG	0.821	0.826	0.830	0.837	0.732	0.952		1.000												
WSDB3	0.995	0.996	0.988	0.978	0.979	0.728	0.930	0.838	1.000											
WSDB4	0.989	0.992	0.988	0.984	0.984	0.829	0.944	0.819	0.994	1.000										
WSDB5	0.975	0.982	0.985	0.986	0.988	0.724	0.889	0.790	0.981	0.986	1.000									
WSDB6	0.960	0.972	0.974	0.987	0.939	0.538	0.697	0.717	0.967	0.974	0.983	1.000								
WSDB7	0.950	0.942	0.964	0.977	0.938	0.456	0.920	0.659	0.957	0.957	0.973	0.982	1.000							
WSDB8	0.949	0.973	0.972	0.958	0.953	0.998	1.000	0.868	0.950	0.954	0.959	0.984	0.970	1.000						
WSDB9	0.983	0.987	0.977	0.989	0.993			0.821	0.984	0.987	0.984	0.994	-0.612	1.000	1.000					
WSDBUGNP	0.857	0.943	0.949	0.923	0.996	0.992	0.957		0.849	0.905	0.943	0.555	0.704			1.000				
WSDBUG	0.866	0.866	0.807	0.774	0.776	0.916	1.000	0.957	0.824	0.777	0.836	0.619	0.741	0.849	0.621		1.000			
WSDC3	0.991	0.992	0.991	0.978	0.941	0.880	0.915	0.653	0.994	0.990	0.984	0.980	0.965	0.933	0.953	0.780	0.792	1.000		
WSDC4	0.990	0.993	0.993	0.992	0.977	0.970	0.995	0.739	0.994	0.992	0.994	0.989	0.977	0.925	0.980	0.856	0.709	0.994	1.000	
WSDC5	0.973	0.983	0.979	0.993	0.985	0.918	0.970	0.680	0.980	0.989	0.989	0.979	0.990	0.972	0.990	0.998	0.615	0.985	0.995	1.000

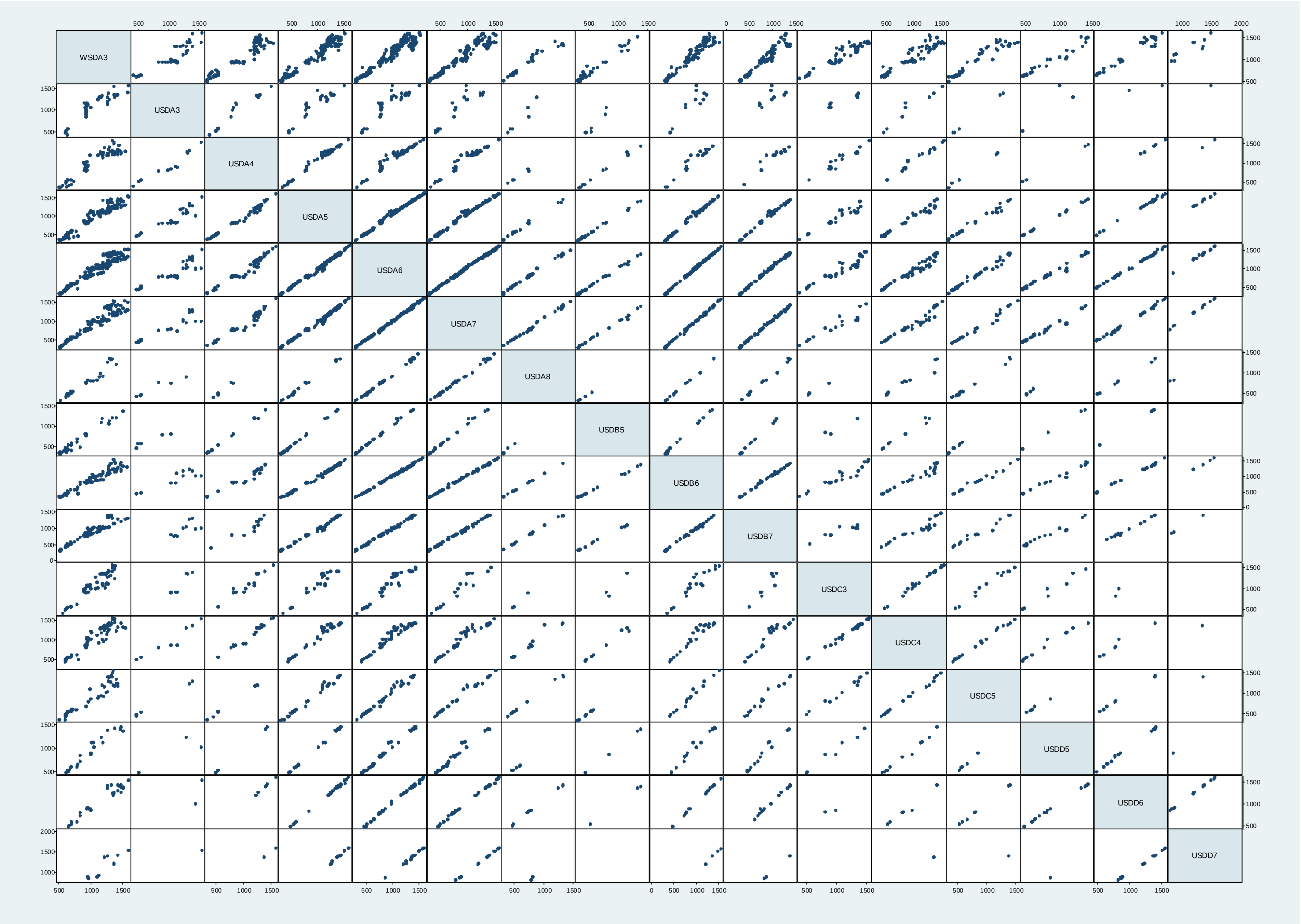


Figure 3.5: WSDA3 and unwashed coffee contracts correlation matrix

Table 3.4: Correlation coefficient matrix WSDA3 with other unwashed Sidama coffee contracts

	WSDA3	*USDA3*	*USDA4*	*USDA5*	*USDA6*	*USDA7*	*USDA8*	*USDB4*	*USDB5*	*USDB6*	*USDB7*	*USDC3*	*USDC4*	*USDC5*	*USDC6*	*USDC7*	*USDD7*
WSDA3	1.000																
USDA3	0.920	1.000															
USDA4	0.952	0.965	1.000														
USDA5	0.973	0.910	0.991	1.000													
USDA6	0.970	0.903	0.977	0.997	1.000												
USDA7	0.969	0.921	0.978	0.996	0.999	1.000											
USDA8	0.980	0.964	0.990	0.998	0.998	0.999	1.000										
USDB4	0.930	0.907	0.992	0.997	0.997	0.998	-	1.000									
USDB5	0.984	0.971	0.998	1.000	0.999	0.998	0.994	0.989	1.000								
USDB6	0.973	0.902	0.964	0.997	1.000	0.999	0.999	0.994	1.000	1.000							
USDB7	0.960	0.672	0.941	0.992	0.999	0.999	0.998	1.000	0.999	0.998	1.000						
USDC3	0.945	0.952	0.952	0.948	0.942	0.943	0.997	-	0.960	0.953	0.919	1.000					
USDC4	0.947	0.965	0.989	0.976	0.962	0.969	0.975	-	0.992	0.969	0.930	0.994	1.000				
USDC5	0.961	1.000	0.999	0.994	0.993	0.993	0.998	1.000	0.995	0.989	0.987	0.989	0.998	1.000			
USDC6	0.975	0.951	0.998	0.998	0.999	0.999	1.000	-	1.000	0.995	0.998	0.974	0.999	1.000	1.000		
USDC7	0.903	-	0.822	0.991	0.993	0.996	-	-	-	0.974	0.990	0.842	0.989	0.999	0.989	1.000	
USDD5	0.983	0.886	1.000	1.000	0.989	0.981	0.989	-	1.000	0.985	0.986	0.982	0.991	0.994	0.988	-	
USDD6	0.969	1.000	0.996	0.998	0.999	0.999	0.999	-	1.000	0.999	0.996	1.000	0.984	1.000	1.000	1.000	
USDD7	0.878	-	1.000	0.994	0.996	1.000	1.000	-	-	0.999	1.000	-	-	-	1.000	0.997	1.000

Scenario II: The relationship between WSDA3 and unwashed Sidama coffee categories

Figure 3.5 and Table 3.4 results indicate that WSDA3 has strong and positive relationship with all unwashed Sidama coffee contracts. Now we have answered the second scenario of the first research question of the study.

Scenario III: Relationship between different coffee origins

The correlation coefficient and scatter plots of WSDA3 with other origin coffee contracts are presented in Annex 4 to Annex 7. The statistical result suggested that WDSA3 has relationship with most of the contracts and it will be used to predict other contracts price.

3.2. Data Preparation for Analysis

A time series database consists of a sequence of values obtained over a repeated measurement of time. The values are typically measured at equal time intervals (hourly, daily, weekly). Time-series databases are popular in many applications, such as stock market analysis, economic and sales forecasting, budgetary analysis, utility studies, inventory studies, yield projections, workload projections, process and quality control, scientific and engineering experiments, and medical treatments [22]. The research's data frequency is daily trade data of coffee contracts.

The major attributes that are employed in the prediction process are the daily closing price of the reference and the predicted contract where the two contracts were traded or co-occur in the same trading date. The other variable is warehouse where graded coffee contracts are stored.

Daily price data

Pi - The price of the reference contract, in this case WSDA3

P i - The price of the predicted contract

Warehouse

One coffee contract might be stored in one or more warehouses located in the country (Table 1.2). In a given trading day same coffee contracts from different warehouse location are sold with different price. Usually, contracts from far warehouse locations are sold with a lesser price. During coffee price quotation the buyer considers the contract storage location because the buyer is responsible to cover the transportation cost. The buyer of the product uses the warehouse information as one attribute to decide the daily price of the commodity. Similarly in this study warehouse information is taken as one attribute to determine to predict the daily selling price of coffee contracts.

CHAPTER FOUR

DESIGN OF THE SYSTEM

4.1. Architecture of the System

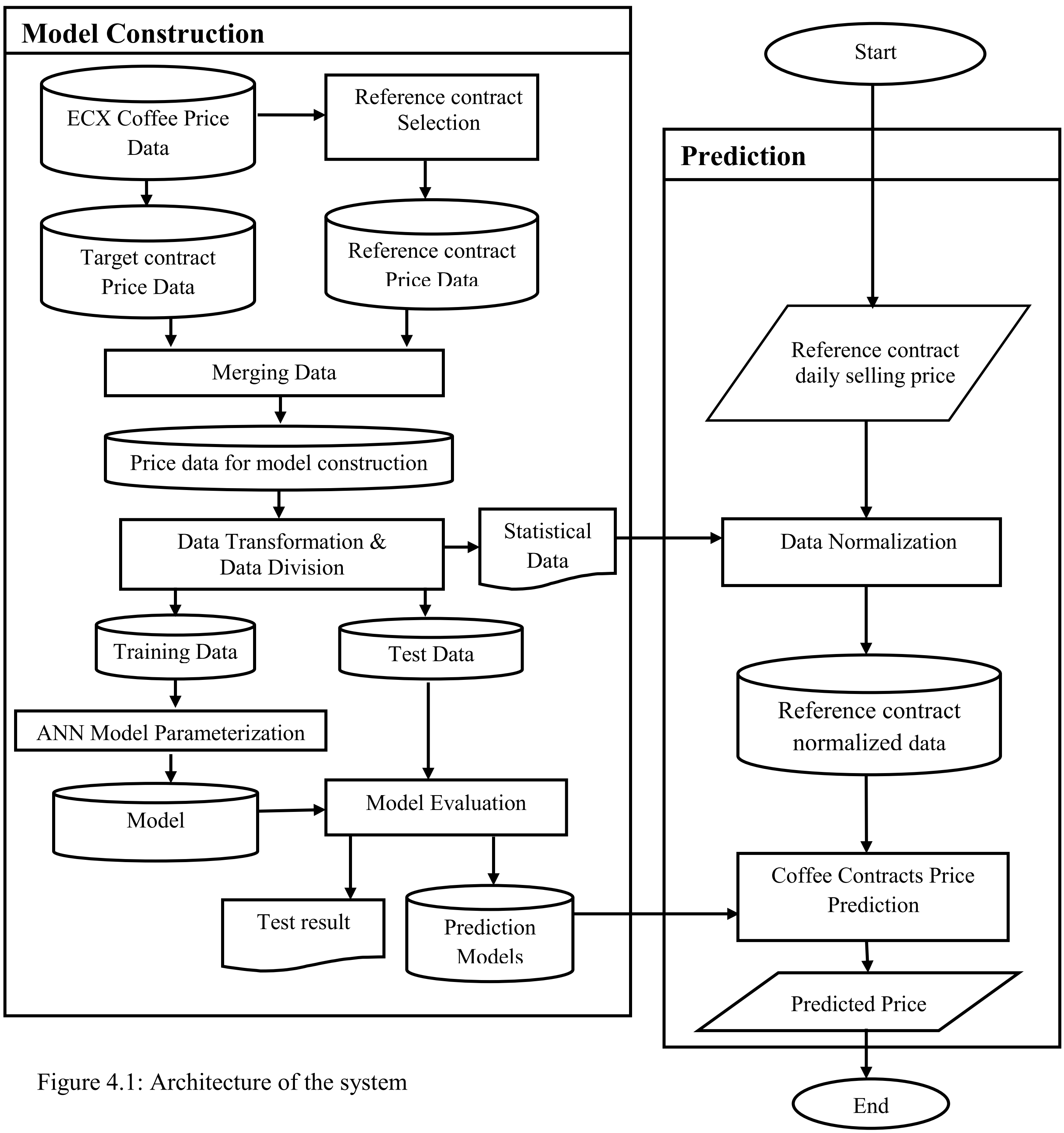


Figure 4.1: Architecture of the system

The architecture of the system model has two major components as depicted in Figure 1.1. The system has two phases the learning phase we call is model construction and the prediction phase we call it prediction.

4.2. Model Construction

The model construction component has four major processed to design the neural network prediction models. They are merging data, data transformation and data division, ANN Model parameterization and model evaluation. Finally, the models are stored in a database called Prediction models as a knowledge base to be used in the prediction component.

4.2.1. Reference Contract Selection

The model construction started from ECX coffee trade data. Detail data analysis was undertaken to select the reference contract (Section 3.1.7). Finally washed Sidama grade A 3 (WSDA3) was selected as a reference contract to predict other coffee contract daily price data. Its daily price data was stored in reference contract price data store.

4.2.2. Merging Data

From ECX coffee trade data the target contract that is going to be predicted from the reference contract is extracted and stored as target contract price data store. To undertake the model construction process the reference contract price data and as target contract price data are merged.

4.2.3. Data Transformation and Data Division

It incorporates two processes the data transformation and data division. The price data for model construction is taken as an input to undertake the transformation and data division processes.

Data transformation is a process of transforming the data into a form which is appropriate for model building task. Neural networks can only process datasets with attributes of numeric value. Besides, the values of each attribute must be between 0 and 1 or between -1 and 1. Therefore, to make the dataset suitable for the neural network, values of all attributes should be changed into numbers between the acceptable ranges. This process is usually known as normalization [53]. Moreover, data transformation is conducted in order to improve the accuracy of the forecast. As most financial prediction models use the hyperbolic tangent function this study also uses the hyperbolic tangent transfer function. In this study data normalization and scaling were used to make the dataset appropriate for the neural network tool.

Normalization of Price Data

There are several techniques used for data normalization which include Z-Score normalization and Min-Max normalization. In min-max normalization technique the input and output values are rescaled from its original range of value in to a new range of values. Most often the input and output variables are rescaled to lie within a range of 0 to 1 or from -1 to 1. It is capable of preserving the exact relationship that exists in the data and it does not introduce any biase [48].

For this study, the reference and target contracts daily closing price data were normalized into [-1, 1] by min-max normalization technique.

$$\alpha = 2\left(\frac{\mathrm{Pi{-}Pmin}}{Prange}\right) - 1$$

$$\alpha\ = 2\left(\frac{\mathrm{P\ i{-}P\ min}}{P\ range}\right) - 1$$

Where

$\alpha - is\ the$ normalized daily price of $the\ reference\ contract$

$\alpha\ - is\ the$ normalized daily price of $the\ \ target\ contract$

$Pi - is\ the\ daily\ \ price\ of\ the\ reference\ contract\ at\ specific\ trade\ date$

$Pi\ \ - is\ the\ daily\ \ price\ of\ the\ \ target\ contract\ at\ specific\ trade\ date$

$Pmin\ \ - is\ the\ minimum\ price\ of\ the\ reference\ \ contract$

$P\ min - is\ the\ minimum\ price\ \ of\ the\ target\ contract$

$Prange = Pmax - Pmin$

Pmin $and\ P\ min$, $Prange\ and\ P\ range$ values for each coffee contracts including the reference were stored as Statistical Data and they will be used for the normalizing the reference price in the prediction component.

Scaling Warehouse information

Warehouse information in the coffee trading dataset is represented in text as Addis Ababa, Awassa, Dilla and Jimma. The textual warehouse information was changed to numerical representation into a range [-1, 1]. To illustrate how this is undertaken

- In almost all trading days, the reference contract was stored in Awassa and it has one warehouse represented by 1.
- WSDA class contracts (WSDA3, WSDA4, WSDA5 and so on) are stored in Awassa; it has only one warehouse information represented by 1.
- WSDB class contracts in Awassa and Dilla; two warehouse information represented by 1 and -1 respectively.
- WSDC in Awassa and Welayta Sodo; two warehouse information represented by 1 and -1 respectively.

To make it generic, if a contract is stored in one warehouse it is represented by 1; if it is stored in two warehouses the first warehouse is represented by 1 and the second is represented by -1. Finally, if it is stored in n warehouse its scalar representation is calculated by dividing the range [-1, 1] into n equal points. The scalar representation of ECX warehouses are presented in Annex 9. Table 4.1, Table 4.2 and Table 4.3 illustrate how data transformation was performed starting from the original price, scaling the warehouse information to fully transformed data. ULK6 is unwashed Lekemte grade 6 contract.

Table 4.1: ULK6 and WSDA3 daily price and warehouse unprocessed data

Trade date	Contract/Warehouse		
	ULK6		WDSA3
	Addis Ababa	Gambela	Awassa
11/19/2010	900	890	950
8/9/2012	916	911	1,030
8/28/2012	930	910	1,010
9/4/2012	900	886	1,020
9/5/2012	890	880	1,050
9/13/2012	940	923	1,140
9/17/2012	960	942	1,200
9/19/2012	955	930	1,180
9/20/2012	950	926	1,160
10/9/2012	970	953	1,100
10/17/2012	975	960	1,060
11/1/2012	935	917	960
11/2/2012	930	910	960
11/5/2012	915	895	1,000
12/6/2012	822	795	985
12/7/2012	825	785	1,010
12/25/2013	940	940	890

Where Addis Ababa is represented by 1.

Gambela is represented by -1.

Table 4.2: ULK6 and WSDA3 daily price and scalled warehouse information

Trade date	Contract/Warehouse		
	ULK6		WDSA3
	Price	warehouse	Price
11/19/2010	900	1	950
11/19/2010	890	-1	950
8/9/2012	916	1	1,030
8/9/2012	911	-1	1,030
8/28/2012	930	1	1,010
8/28/2012	910	-1	1,010
9/4/2012	900	1	1,020
9/4/2012	886	-1	1,020
9/5/2012	890	1	1,050
9/5/2012	880	-1	1,050
9/13/2012	940	1	1,140
9/13/2012	923	-1	1,140
9/17/2012	960	1	1,200
9/17/2012	942	-1	1,200
9/19/2012	955	1	1,180
9/19/2012	930	-1	1,180
9/20/2012	950	1	1,160
9/20/2012	926	-1	1,160
10/9/2012	970	1	1,100
10/9/2012	953	-1	1,100
10/17/2012	975	1	1,060
10/17/2012	960	-1	1,060
11/1/2012	935	1	960
11/1/2012	917	-1	960

Table 4.3: The transformed price data of ULK6 and WSDA

Trade date	Input		Target/ Desired output
	WSDA3	WH	ULK6
11/19/2010	-0.61290	1	0.21053
11/19/2010	-0.61290	-1	0.10526
8/9/2012	-0.09677	1	0.37895
8/9/2012	-0.09677	-1	0.32632
8/28/2012	-0.22581	1	0.52632
8/28/2012	-0.22581	-1	0.31579
9/4/2012	-0.16129	1	0.21053
9/4/2012	-0.16129	-1	0.06316
9/5/2012	0.03226	1	0.10526
9/5/2012	0.03226	-1	0.00000
9/13/2012	0.61290	1	0.63158
9/13/2012	0.61290	-1	0.45263
9/17/2012	1.00000	1	0.84211
9/17/2012	1.00000	-1	0.65263
9/19/2012	0.87097	1	0.78947
9/19/2012	0.87097	-1	0.52632
9/20/2012	0.74194	1	0.73684
9/20/2012	0.74194	-1	0.48421
10/9/2012	0.35484	1	0.94737
10/9/2012	0.35484	-1	0.76842
10/17/2012	0.09677	1	1.00000
10/17/2012	0.09677	-1	0.84211
11/1/2012	-0.54839	1	0.57895
11/1/2012	-0.54839	-1	0.38947
11/2/2012	-0.54839	1	0.52632
11/2/2012	-0.54839	-1	0.31579

Data Division

The transformed data is taken as an input of this process. It divides into the training data set and test data. The training data sets are used to learn the pattern present in the data and the test set is used to evaluate the generalization ability of the fitted model. Data are divided by random selection using the Dividerand[7] [40]. In this study, 70% of the data set is used for training and they are store as training data the remaining 30% are used to test the goodness of the model and stored as test data.

[7] Dividerand:- Matlab function divide the target dataset using random indices [40].

4.2.4. ANN Model Parameterization

To create coffee daily price prediction models the two feed forward neural network models the MLP and RBF were used. This process is used to select a model and to set parameters for each model.

4.2.5. Multilayer Preceptor Neural Network

The choice of the network architecture and the choice of the parameters are the crucial factor to build an accurate MLP prediction model. The performance of neural networks depends on the network type, number of hidden layers, and number of nodes in the hidden layer, the learning algorithm and the activation function used at each layer of the network [31, 33]. The selection of these parameters will be discussed in this section.

Number of input nodes

Success in designing a neural network depends on clear understanding of the problem; identifying the relevant attributes for the prediction process. The number of neurons in the input layer is the easiest parameter to select since it is defined automatically according to the number of inputs for the ANN [33]. In this study, the input neurons are the normalized daily price of the reference contract and warehouses of the predicted contract.

Number of hidden layers

The hidden layer(s) allow the network to capture and generalize the pattern in the data and allows the network to generate numerous relationships between the inputs and the outputs. In theory, a neural network with one hidden layer with sufficient number of hidden nodes is capable of approximating any continuous function. In practice, neural networks with one or occasionally two hidden layers are widely used and have performed very well in many studies. Increasing the number of hidden layers also increases the computational time and the danger of over fitting which leads to poor out sample performance. Over fitting occurs when a forecasting model has too few degree of freedom, i.e. when it has a relatively few observation than the parameters and therefore it is able to memorize individual data points than learn the general pattern of the data [31, 33].

It is common practice in the field of financial forecasting to use only one layer and very occasionally two layers. Both theory and the majority of empirical work suggest that ANNs with more than two hidden layers will not improve the networks performance [33]. A network with only one hidden layer has shown to perform very well [55]. Mostly, neural networks with single hidden layer have good universal approximation ability and relatively easier to train.

To determine the number of neurons in the hidden layer, we have used trial and error by constructive approach starting from least number of neuron then keep increasing to improve the network performance. In this research, one and two hidden layer networks were used to build the models and their performance were analyzed and one hidden layer neuron results in best performance and it will be used to build the models.

Number of Neuron in the Hidden Layer

The number of neurons in the hidden layer determines not only the network complexity to model non-linear and interactive behavior of the data but also the networks ability to learn and generalize. Too many or too few neurons in the hidden layer will cause the over fitting or under fitting problem.

The best approach to find the optimal number of hidden units is trial and error. Selecting the optimal number of neurons in the hidden layer through experimentations through three different methods, namely through a fixed, constructive, and destructive approach. The fixed approach trains a group of ANNs separately with different numbers of hidden neurons to find the optimal network. The constructive and destructive approaches involve changing the number of hidden neurons during the training by either adding or removing neurons respectively [33].

Generally speaking it is better to have too many hidden units than too few. With too few hidden units, the model might not have enough flexibility to capture the nonlinearities in the data; with too many hidden units, the extra weights can be shrunk toward zero if appropriate regularization is used. Typically, the number of hidden units is somewhere in the range of 5 to 100 [26].

In this research, experimentations are conducted through the constructive approach. We started the experiments with five neurons in the hidden layer for the first training and the performance is evaluated and then the experiment continues by adding neurons in hidden layers until optimum result is obtained by trial and error. And we have observed that the performance of MLP network was optimal when 10 neurons in the hidden layer were used and 10 neurons in the hidden layer were used for all MLP experiments.

Number of Output Nodes

Determining the number of output neuron is straight forward since there is a compelling reason to always use only one output neuron. Kaastra and Boyd [33] noted that the ANN with multiple outputs, especially where the multiple outputs have a wide range would produce lower quality results compared to an ANN with a single output. And thus it is recommended to always use one output neuron. In this study, the normalized daily price of the predicted coffee contract is used as the target.

Transfer Functions

Transfer functions are mathematical formulas that determine the output of the processing neuron. The purpose of transfer function is to prevent outputs from reaching very large values that can thereby inhibit the training process [33]. The choice of transfer functions may strongly influence complexity and performance of neural networks.

In neural network applications the input data is usually scaled into the range of [0, 1] or [-1, 1] according to the activation function of the neurons. Many stock market data is normalized into the range of [-1, 1] and the neural networks are trained and tested using the back propagation algorithm [43]. The range of Tanh is within [-1, 1].

There are many transfer functions some of them are the sigmoid function, the hyperbolic tangent (Tanh activation functions), step, ramping, acr, tan, linear. The sigmoid function and the hyperbolic tangent function are the most used transfer functions in time series prediction [55].

Bishop [3] uses the hyperbolic tangent function which is more extensible than the conventional sigmoid transfer function. This transfer function often gives rise to faster convergence of the training network than the logistic function.

Karlik and Olgac [34] studied the performance of five activation functions of MLP architecture and compare the performance activation functions Bi-polar sigmoid, Uni-polar sigmoid, Tanh, and Conic Section. The neural network computed good results when “Tanh-Tanh” combination of activation functions was used for both neurons of hidden and output layers.

For this study, hyperbolic transfer function is used in the hidden and output layers because it is more extensible and converges faster than other method. Moreover, it is widely applicable for prediction problems and resulted in good prediction result.

Training Function

The goal of any training algorithm is to find the logical relationship from the given input/output and to minimize the error between the actual output data and the results obtained from the ANN. There are many different learning algorithms for different architectures.

In ANN, the learning methods are broadly categorized as supervised and unsupervised [26]. In supervised training, initially the network is provided a set of input and outputs and trained. The network then processes the inputs and compares its resulting outputs against the desired outputs. During the training process, the weights and biases of the network are iteratively adjusted to minimize the error calculated by the network performance function. While in unsupervised training, the network is provided with inputs but not with desired outputs. The system decides what features to use to group the input data [48].

Back propagation algorithm is a general purpose supervised learning algorithm. It consists of two phases: the forward phase where the activations are propagated from the input to the output layer, and the backward phase, where the error between the actual and the requested nominal value in the output layer are propagated backwards in order to modify the weights and bias values. It is the best algorithm to be used in MLP because it reduces an error between the actual output and desired output in a gradient descent manner. Moreover, it is widely applied to learning model because it is simple and easy to understand. The back propagation algorithm is used for training the neural network with the Tanh activation functions at all layers of the network to generate relationship between ECX coffee contracts.

Finally, a three layered MLP network with 10 neurons in the hidden layer will be used to construct the MLP models (Figure 4.2).

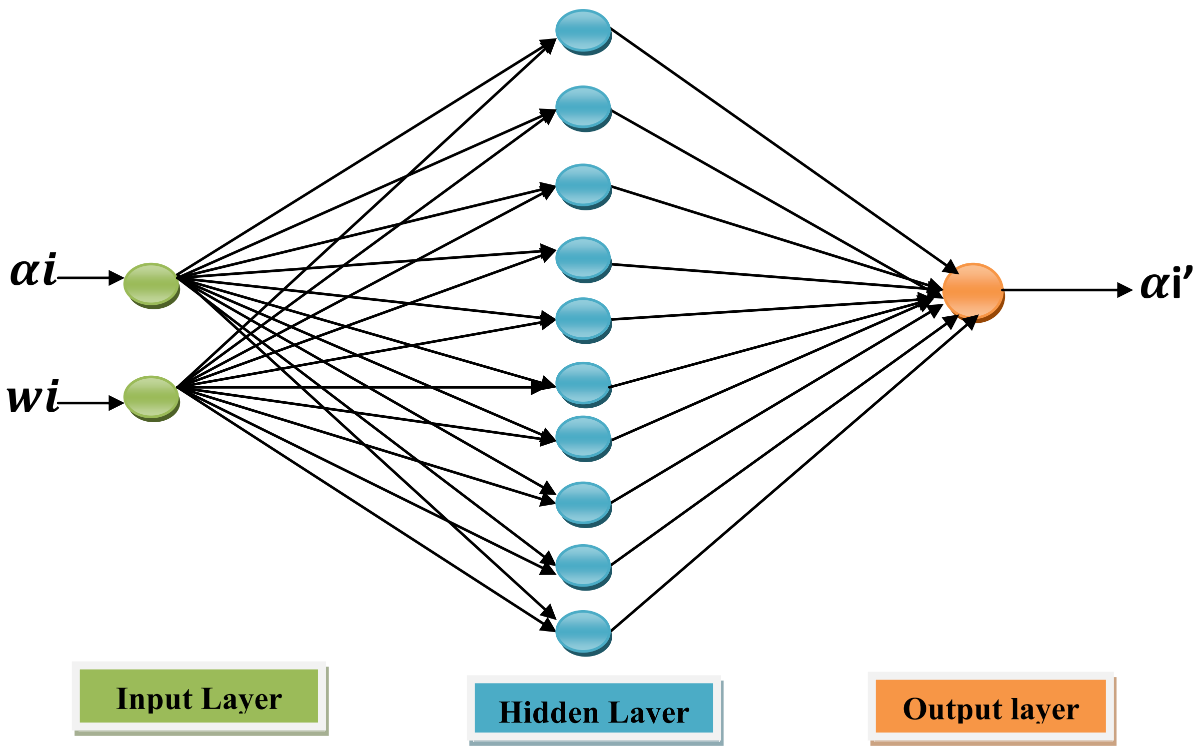


Figure 4.2: MLP neural network architecture for ECX coffee contracts price prediction

In order to get an optimal network structure the researchers have set the following network parameters to build MLP network (Table 4.4).

Table 4.4: MLP parameter set for the experiment

Parameter	Values
Input neurons	2
Network layer	3
Number of hidden layer neurons	10
Transfer function	hyperbolic tangent function; input to hidden layer and from the hidden layer to the output layer
Training function	Trainlm
Output neuron	1
maximum epochs (iterations)	1,000
Error goal	0
Evaluation metrics	R-Squared and MSE

4.2.6. Radial basis Function

The second method is radial basis function neural network. The input and output neurons of RBF network is the same as that of MLP network. The number of neurons in the hidden layer has automatically increased until the precision meets the requirement or the number of neurons reaches the maximum.

The major parameter of the RBF networks is the spread. To find the best RBF model, many experiments were realized by changing the spread value from 0.0001 to 1. After comparing the results of the models we have found that the optimal network is resulted when the spread value is 0.001 and this spread value of 0.001 is used for all RBF experiments.

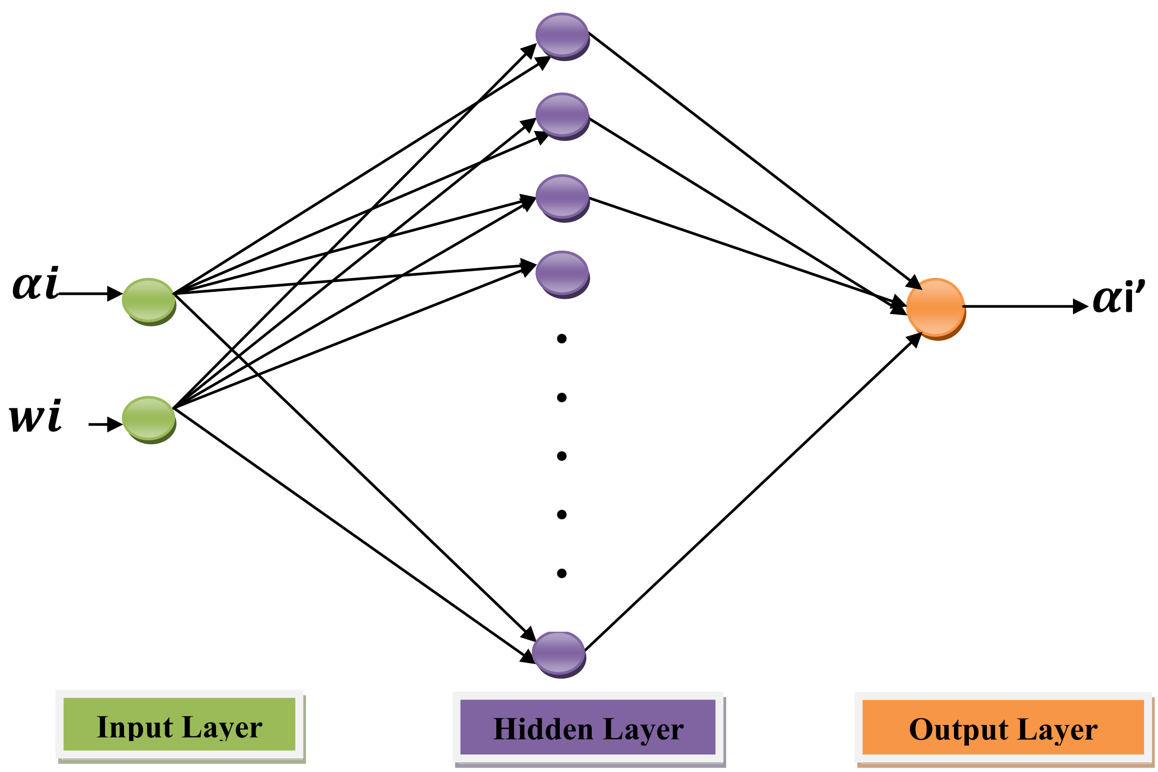


Figure 4.3: RBF neural network architecture for ECX coffee contracts price prediction

Figure 4.3 is the RBF neural network architecture to be used for ECX coffee contracts price prediction. The Gaussians transfer function was used in the hidden layer and linear transfer function in the output layer.

4.2.7. Model Evaluation

After the prediction models are built the next process is to evaluate the goodness of the models. In this study, the performances of the models are evaluated by coefficient of determination (R^2) and Mean Squared Error (MSE). In almost all prediction problems the performance of a model is measured by R^2 and MSE [2, 5, 12, 54].

The coefficient of determination

The coefficient of determination (R^2) provides a measure of how well future samples are likely to be predicted by the model. R^2 describes the proportion of variance of the dependent variable explained by the model [39].

If p' is the predicted value of the i-th sample and p is the corresponding true value, then the R^2 is computed as:-

$$R^2 = \frac{SSR}{SST} \qquad \text{OR} \qquad R^2 = 1 - \frac{SSE}{SST}$$

Where

- SST = total sum of squares $SST = \sum (p - \bar{p})^2$
  - ➢ Measures the variation of the p_i values around their mean p.
- SSE = error sum of squares $SSE = \sum (p - p')^2$
  - ➢ Variation attributable to factors other than the relationship between p and p'.
- SSR = regression sum of squares $SSR = \sum (p' - \bar{p})^2$
  - ➢ Explained variation attributable to the relationship between p and p'.

R-squared value ranges from zero to one, with zero indicating a random relationship and one indicating perfect prediction; higher R-squared values indicate better prediction. In most study R-squared value greater than 0.8 is taken as good accuracy result.

Mean Squared Error

Mean Squared Error (MSE) is the average of the squared values of the prediction errors or squared values of the deviations from the target [55]. It provides information on the average squared error between the network outputs and the target. Generally, MSE can evaluate and compare the predictive power of the models; how close are the predicted values to the target. The lower the MSE, the more accurate is the estimation, zero MSE means no error.

The mean squared error (MSE) estimated over n-samples is defined as

$$MSE(\mathrm{p}, \mathrm{p}') = \frac{1}{\mathrm{nsamples}} + \sum_{n=0}^{\mathrm{nsamples}} (\mathrm{pi} - \mathrm{pi}')^2$$

The constructed models are tested using the test dataset. The models with large R-squared values and small Mean Squared Error are selected and stored as a prediction model knowledge base to predict the selling price of export coffee contracts from the reference contract. The prediction model incorporates the models for each coffee contract along with its warehouse information.

4.3. Prediction

The major processes that are undertaken to predict the daily selling price of export coffee contracts are data normalization followed by coffee contracts price prediction.

4.3.1. Data Normalization

This process accepts the reference contract price data and it takes the statistical data which is the result of the transformation process and it normalizes reference contract price data accepted from the used into the range [-1, 1] by the min-max normalization method (Section 4.2.3). The normalized data will be passed to the price prediction process to predict the daily selling price of all export coffee contracts.

4.3.2. Coffee Contracts Price Prediction

The coffee contracts price prediction is a process designed to predict the daily selling price of all export coffee contract. The input of this prediction process is the normalized reference contract data. This process obtains the models for each coffee contract along with its warehouse information from the prediction models after that prediction will take place. The output of this prediction process is the normalized price data of the target contracts.

To obtain the actual predicted price data the normalized price data need to be de-normalized. De-normalization is a sub-process in the coffee contracts price prediction which converts normalized price data into actual price value.

The de-normalization formula is given below where it is described in section 4.2.3.

$$Pi \ = 2\left(\frac{\mathrm{i+1}}{2}\right) + \ \mathrm{Pmin}$$

Finally, the predicted price of export coffee contracts will be displayed.

.

CHAPTER FIVE
EXPERIMENT

This Chapter presents the data set used in the experiment, the experimental setup, experimental result and discussion of result.

5.1. Data Set

ECX historical coffee trade data from 1/1/2009 to 30/6/2014, about 1,347 trading day data was gathered from ECX database. Relevant attributes are selected from the whole dataset and the datasets are preprocessed to be utilized by the neural network tool. The transformed data was divided into two subsets training data set and test data set. In this study, the random data division method was adopted and 70% of the data set is used for training and 30% for testing the goodness of the model.

5.2. Experimental Setup

The two commonly used feed forward neural network methods will be employed to conduct the experiment. The multilayer perceptron neural network and the radial basis function neural network methods are selected because they are used in many prediction problems and many research works have been reported as they are efficient for function approximation problems (Section 2.8).

To test and validate the models, two statistical metrics; the correlation of determination and the mean squared error were used. A model with large R^2 value on the testing data and small MSE will be selected. To improve the performance of the models the models were trained repeatedly until the best performance was obtained.

Based on the detailed research objective of the study we have designed three experiments:-

Experiment 1:- Washed Sidama Coffee Contracts Prediction from WSDA3.

Experiment 2:- Unwashed Sidama Coffee Contracts Prediction from WSDA3.

Experiment 3:- Other Origin Coffee Contracts Prediction from WSDA3

The experiments are done with Matlab Software (version R2013a) on a computer that is running with Windows 7 and it has 4-core i7-processor running at 2.2 GHz and 4 GB memory.

5.3. Result and Discussion

In this Section, analysis and discussion of experimental results are presented.

5.3.1. Experiment 1: Washed Sidama Coffee Contracts Prediction

Experiment 1 is designed to predict the daily selling price of washed Sidama coffee contracts from WSDA3.

Predicting WSDA4 (one contract from of washed Sidama coffee contracts) from WSDA3

The sample daily coffee price data and transformed daily coffee price of WSDA4 and WSDA3 contracts are presented in the Table 5.1 Columns 5 and 6 are input neurons and Column 7 is the output neurons. To predict WSDA4 from WSDA3 we need to transform the price data by Min-Max approach and the warehouse information to its scalar representation as follows.

Table 5.1: Sample data for WSDA4 and WSDA3 contracts

trade day	Original data			Transformed data		
	WSDA3 Price	Warehouse	WSDA4 Price	WSDA3	Warehouse	WSDA4
1	2	3	4	5	6	7
11/19/10	950	Awassa	900	-0.64103	1	-0.65131
11/24/10	963	Awassa	935	-0.60769	1	-0.56413
11/26/10	966	Awassa	943	-0.60000	1	-0.54421
11/29/10	975	Awassa	953	-0.57692	1	-0.51930
11/30/10	980	Awassa	960	-0.56410	1	-0.50187
12/01/10	996	Awassa	980	-0.52308	1	-0.45205
12/03/10	1,015	Awassa	1,000	-0.47436	1	-0.40224
12/06/10	1,060	Awassa	1,045	-0.35897	1	-0.29016
12/07/10	1,100	Awassa	1,075	-0.25641	1	-0.21544
12/08/10	1,155	Awassa	1,129	-0.11538	1	-0.08095
12/10/10	1,240	Awassa	1,210	0.10256	1	0.12080
12/13/10	1,220	Awassa	1,200	0.05128	1	0.09589
12/14/10	1,220	Awassa	1,180	0.05128	1	0.04608
12/15/10	1,185	Awassa	1,170	-0.03846	1	0.02117
12/16/10	1,165	Awassa	1,165	-0.08974	1	0.00872
12/20/10	1,135	Awassa	1,118	-0.16667	1	-0.10834
12/21/10	1,130	Awassa	1,110	-0.17949	1	-0.12827
12/22/10	1,184	Awassa	1,166	-0.04103	1	0.01121

trade day	Original data			Transformed data		
	WSDA3 Price	Warehouse	WSDA4 Price	WSDA3	Warehouse	WSDA4
12/23/10	1,244	Awassa	1,225	0.11282	1	0.15816
12/24/10	1,240	Awassa	1,225	0.10256	1	0.15816
12/27/10	1,215	Awassa	1,200	0.03846	1	0.09589
12/28/10	1,215	Awassa	1,200	0.03846	1	0.09589

Models were built by MLP and RBF methods. The MLP model resulted in R^2 value of 99.41% and MSE of 0.0028 while RBF resulted in R^2 of 66.04% and MSE of 0.0760.

WSDA4 actual test data along with simulated results by MLP and RBF methods are presented in Annex 10. Figure 5.1 depicts comparison between simulated results of the two methods and actual price data on the test dataset.

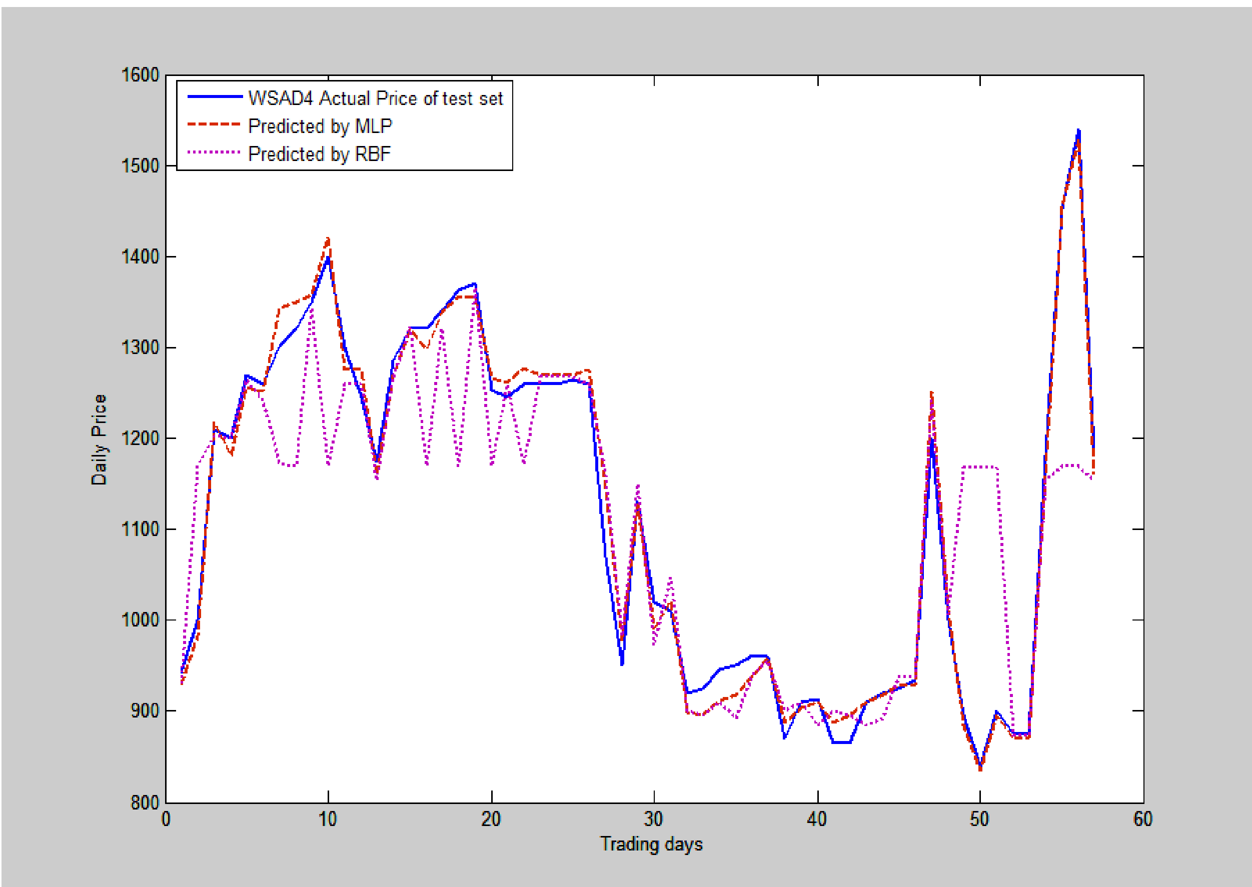


Figure 5.1: WSDA4 Actual test data against simulated results of MLP and RBF Methods

The WSDA4 price test data sets are closely fitted by MLP model but RBF poorly approximate the test data set. From the statistical and graphical results we have concluded that MLP function approximation methods are able to capture the relationship between the references and WSDA4 and generate good prediction results.

Similar to WSDA4 all washed Sidama coffee contracts prediction models were developed a total of 26 experiments were conducts and each experiment was trained many times until good accuracy and MSE result was found. The experimental results are presented in the Table 5.2.

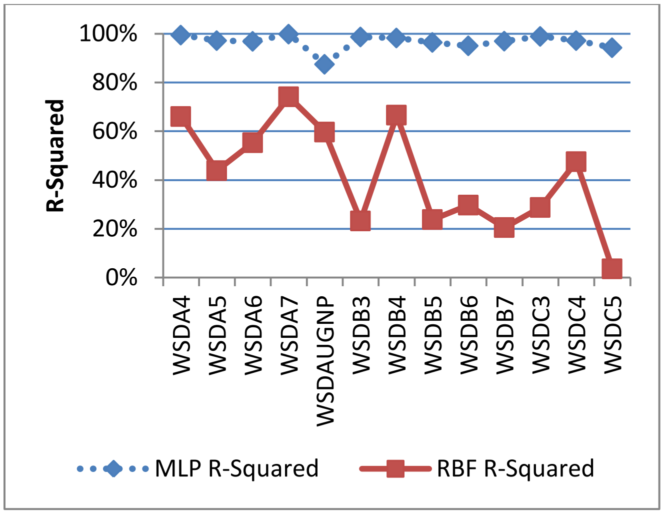
Table 5.2: Sidama washed coffee contracts prediction result

S.N.	**Coffee Contract**	**Warehouse**	**Trading days**[8]	**MLP**		**RBF**	
				R^2	**MSE**	**R^2**	**MSE**
1	WSDA4	AW=1	38	99.41%	0.0028	66.04%	0.0760
2	WSDA5	AW=1	176	97.14%	0.0121	43.86%	0.1333
3	WSDA6	AW=1	176	96.84%	0.0192	55.31%	0.1011
4	WSDA7	AW=1	32	99.80%	0.0193	74.14%	0.6199
5	WSDAUGNP	AW=1	65	87.45%	0.1047	59.68%	0.0926
6	WSDB3	AW= 1; DL= -1	374	98.64%	0.0072	23.25%	0.2018
7	WSDB4	AW= 1; DL= -1	468	98.22%	0.0099	66.62%	0.0939
8	WSDB5	AW= 1; DL= -1	101	96.39%	0.0290	23.83%	0.3129
9	WSDB6	AW= 1; DL= -1	85	95.01%	0.0367	29.75%	0.2320
10	WSDB7	AW= 1; DL= -1	29	96.92%	0.0512	20.54%	0.1847
11	WSDC3	AW= 1; WS= -1	176	98.86%	0.0093	28.77%	0.2794
12	WSDC4	AW= 1; WS= -1	95	97.13%	0.0140	47.54%	0.1225
13	WSDC5	AW= 1; WS= -1	55	94.23%	0.0201	3.62%	0.1169

From the above table, MLP predictive model provides better accuracy results than RBF for all Sidama washed coffee contracts with large value and small MSE.

The performances of MLP and RBF methods on each contract were compared and presented in Figure 5.2 and Figure 5.3.

[8] Trading days means the number of days the two contracts co-occur together



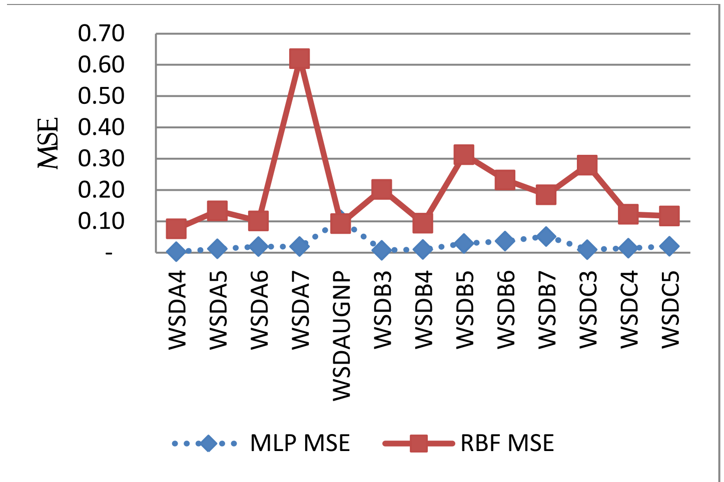


Figure 5.2: Analysis of Experiment_1 Using R2

Figure 5.3: Analysis of Experiment_1 using MSE

Comparing the empirical results of MLP with RBF neural network; MLP based models perform near 100% accuracy (R-square values) whereas RBF accuracy result fluctuates from 60% to near 3%.

MSE values of MLP models are smaller than that of RBF model. We have found that MLP performs better prediction than RBF. Washed Sidama coffee contracts can be predicted by MLP model with higher accuracy.

5.3.2. Experiment 2: Unwashed Sidama Coffee Contacts Prediction

The aim of the second experiment is to predict unwashed Sidama coffee contracts from WSDA3. Unwashed Sidama coffee contracts are highly correlated with WSDA3 contract as explained in Section 3.1.8 of the study.

18 experiments were undertaken where each experiment was trained many times until good results are found. The results of this experiment are presented in Table 5.3.

Table 5.3: The prediction of result of unwashed Sidama coffee contracts

S.N.	Contract	Warehouse	Trading Days	MLP		RBF	
				R^2	MSE	R^2	MSE
1	USDA4	AW=1	138	86.34%	0.0935	49.47%	0.1562
2	USDA5	AW=1	277	89.45%	0.0578	54.05%	0.1304
3	USDA6	AW=1	335	88.19%	0.0817	32.69%	0.1924
4	USDA7	AW=1	150	80.69%	0.0918	44.02%	0.1437
5	USDB5	DL=1	44	94.10%	0.0681	67.05%	0.0681
6	USDB6	AW= 1; DL= -1	114	91.48%	0.0610	25.41%	0.2739
7	USDB7	AW= 1 ;DL= -1	122	88.29%	0.0644	19.49%	0.3251
8	USDC3	WS=1	133	89.58%	0.0791	45.14%	0.2097
9	USDC4	WS=1	91	81.62%	0.1171	0.45%	0.3148

Generally experimental result shows that MLP method resulted in higher accuracy than RBF.

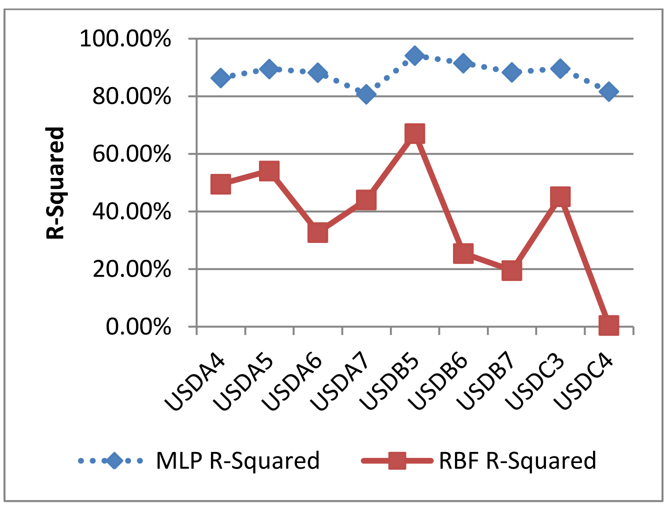


Figure 5.4: Analysis of Experiment_2 using R^2

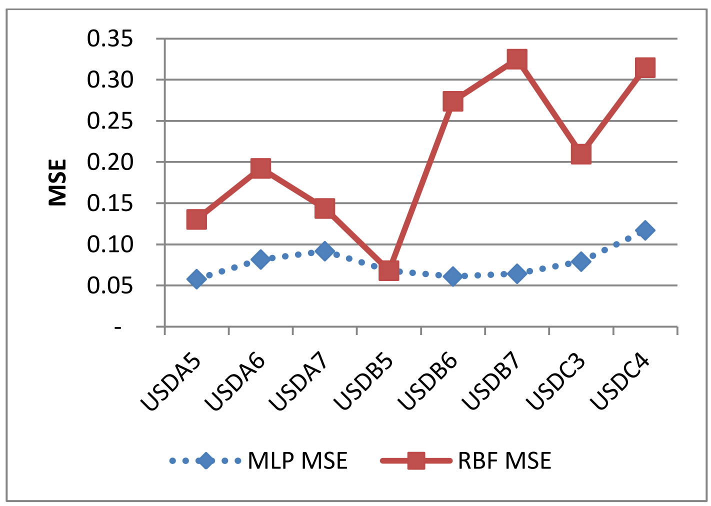


Figure 5.5: Analysis of Experiment_2 using MSE

MLP based model result in high accuracy than RBF models for all unwashed Sidama coffee contracts. Moreover, MSE of MLP is smaller than that of RBF. From the experimental results (Figure 5.4 and Figure 5.5), it was observed that the multilayer perceptron network was able to predict the daily selling price of all unwashed coffee contracts with higher accuracy.

5.3.3. Experiment 3: Other Origins Coffee Contracts Prediction

Ethiopia is the producer for the several well-known varieties of coffee including Yirgachefe, Jimma, Tepi, Lekemete, Harar, Limmu, Bebeka, Forest, Kafa and bale. Experiment 3 was subdivided into two based on coffee processing types, washed and unwashed of different origins.

Experiment 3.1–Washed Coffee of different origin

In this experiment washed coffee contracts of different origins were predicted from the reference contract. 30 experiments each were trained many times until good results are accuracy result was found. The performances of the models are depicted in Table 5.4.

Table 5.4: Washed Coffee of different origins

S.N.	Origin	Contract	Trading days	MLP		RBF	
				R^2	MSE	R^2	MSE
1	**Limmu**	WLMA3	236	97.35%	0.0165	66.93%	0.1021
2		WLMA4	179	98.02%	0.0198	42.37%	0.2760
3		WLMA5	73	86.54%	0.1266	44.09%	0.2322
4		WLMB3	44	98.70%	0.0109	55.18%	0.1862
5		WLMB4	63	96.24%	0.0166	50.54%	0.0868
6		WLMB5	34	97.71%	0.1354	42.42%	0.2611
7	**Bebeka**	WBB3	98	98.92%	0.0146	38.05%	0.1321
8		WBB4	70	89.62%	0.0684	16.50%	0.1994
9	**Yirgachefe**	WYCA3	295	91.23%	0.0396	30.28%	0.1141
10		WYCA4	120	93.45%	0.0291	26.56%	0.1404
11		WYCB3	239	96.00%	0.0171	36.45%	0.1386
12		WYCB4	170	93.03%	0.0270	26.09%	0.1275
13	**Lekemte**	WLK3	84	77.22%	0.1402	19.33%	0.0240
14		WLK4	84	95.93%	0.0205	74.05%	0.0634
15	**TEPI**	WTP3	91	96.20%	0.0526	2.51%	0.3753

Generally good prediction result was obtained from MLP algorithm. The MLP and RBF method comparative analysis for each washed coffee contracts of other origins are presented in Figure 5.6 and Figure 5.7.

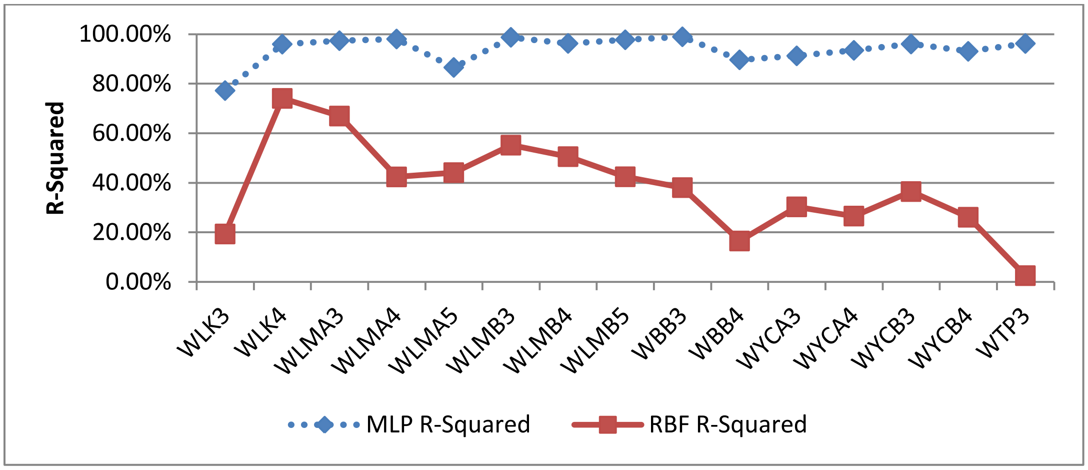


Figure 5.6: Analysis of experiment 3.1 using R^2

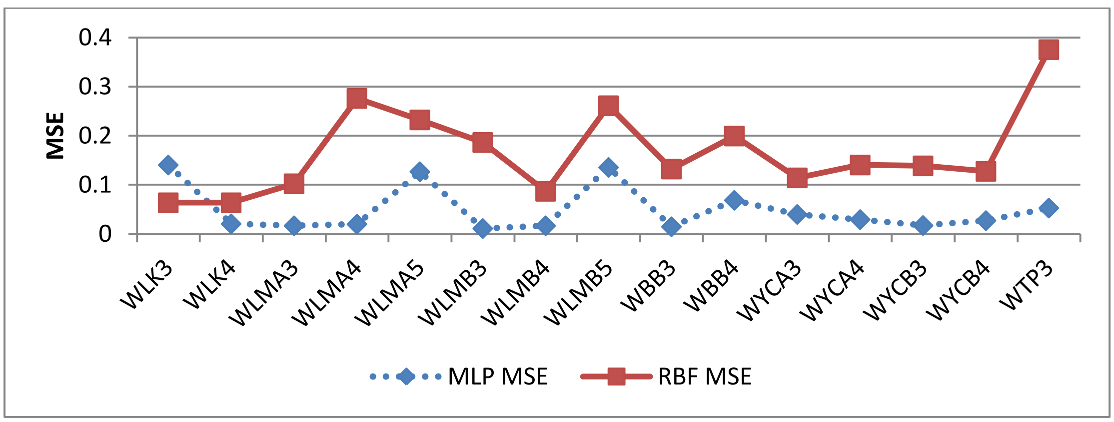


Figure 5.7: Analysis of experiment 3.1 using MSE

From the above graphs MLP r-squared values are larger than RBF for all contracts under this category. Moreover its mean squared error values are smaller than RBF for all except WLK4. MLP based models outperform RBF based models in predicting the daily selling price of washed coffee contracts of different origins from WSDA3.

Experiment 3.2–Unwashed Coffee of different origin

Unwashed coffee contracts of different origins were predicted from the reference contract. Here 30 experiments were conducted where each experiment were trained a number of times to find the optimum network result and the performances of the models are depicted in Table 5.5.

Table 5.5: Unwashed Coffee of different origins

S.N.	Origin	Contract	Trading days	Warehouse	MLP		RBF	
					R^2	MSE	R^2	MSE
1	Lekemte	ULK5	394	GM=1	81.77%	0.0877	51.05%	0.1280
2		ULK6	486	GM = 1; AA = -1	82.38%	0.0746	40.20%	0.1386
3		ULK7	453	GM = 1; AA = -1	92.75%	0.0288	40.14%	0.1073
4		ULK8	320	GM = 1; AA = -1	86.88%	0.0538	29.09%	0.1513
5	Jimma	UJMA7	487	BD=1	87.70%	0.0490	50.09%	0.1047
6		UJMA8	451	JM=1	86.51%	0.0605	49.60%	0.1211
7		UJMB6	408	BD=1	87.53%	0.0513	41.44%	0.1212
8		UJMB7	461	BD=1	84.71%	0.0633	55.74%	0.0948
9		UJMB8	302	BD=1	85.99%	0.0752	44.65%	0.1547
10	Yirgachefe	UYCB5	143	DL=1	86.05%	0.0515	35.54%	0.1156
11		UYCB7	231	DL=1	90.80%	0.0790	33.26%	0.2773
12	FOREST	UFRA7	232	BO = 1	90.06%	0.1984	53.26%	0.0807
13		UFRA8	398	BO = 1; JM = -1	88.89%	0.0533	22.92%	0.1950
14		UFRA9	470	BO = 1; JM = -1	89.85%	0.0406	52.77%	0.9940
15		UFRAUG	506	BO = 1; JM = -1	90.25%	0.0425	47.09%	0.1186

From Table 5.5 MLP performs better than RBF for all contracts.

Let's see this by taking Unwashed Lekemte grade 7(ULK7) is accuracy by MLP is 92.75% and RBF is 40.14%. Figure 5.8 the ULK test data are best approximated by MLP than RBF the data are presented in Annex 11.

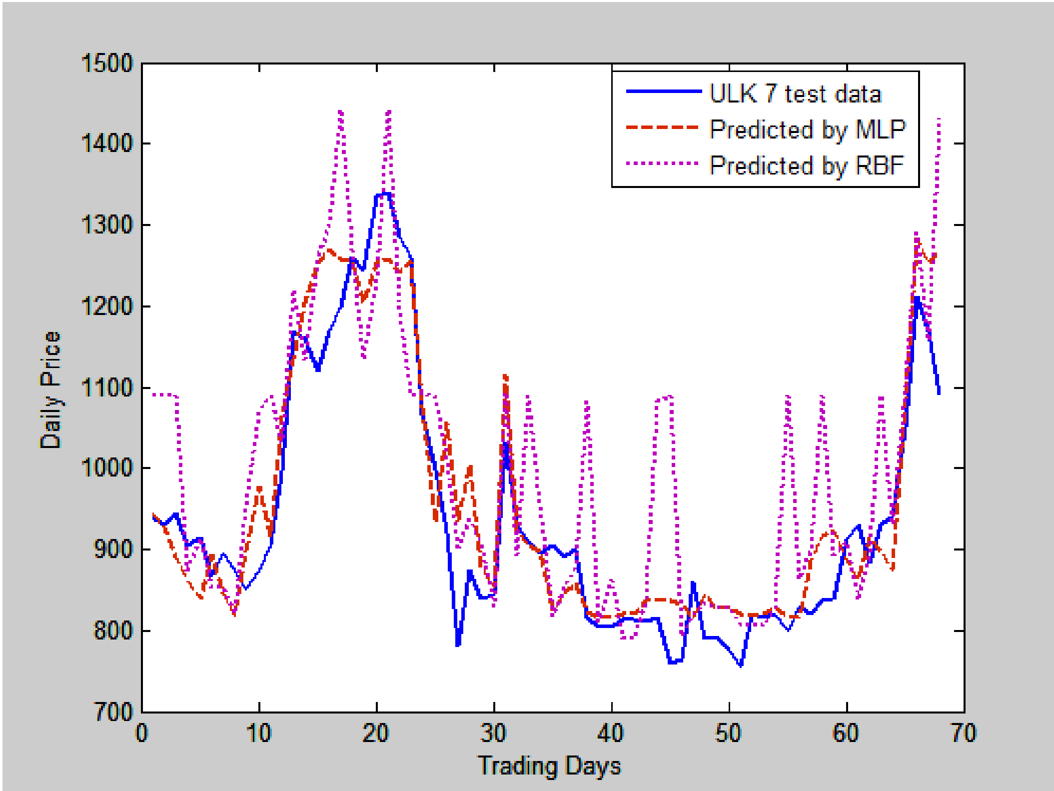


Figure 5.8: Actual test data of ULK6 approximated by MLP and RBF models

5.3.4. Summary of Experimental Result

The experimental results of the study are summarized by major statistics minimum, maximum, average and variance of R2 and MSE values measured in each experiment. The summaries are presented in Table 5.6 and Table 5.7.

Table 5.6: Analysis of R^2

S. N	Experiment	MLP Model				RBF Model			
		Minimum	Maximum	Average	Variance	Minimum	Maximum	Average	Variance
1	Experiment-1	87.45%	99.80%	96.62%	3.20%	3.62%	74.14%	41.77%	21.76%
2	Experiment-2	80.69%	94.10%	87.75%	4.34%	0.45%	67.05%	37.53%	20.18%
3	Experiment-3.1	77.22%	98.92%	93.74%	5.79%	2.51%	74.05%	38.09%	19.03%
4	Experiment-3.2	81.77%	92.75%	87.47%	3.08%	22.92%	55.74%	43.12%	9.66%

R^2 Analysis by Method: - Large R^2 values with small variance were obtained in MLP models. Its minimum value is >77%. On the other hand the smallest R^2 value with relatively larger variance is obtained in RBF based models. And thus MLP result is better accuracy than RBF in all experiments.

R^2 Analysis by Experiment:- In MLP base models the largest R^2 with smallest variance is achieved in Experiment-1. Second best result is observed in Experiment-3.1. Moreover their average R^2 values are 96.62% and 93.74% respectively. Both experiments are undertaken for washed coffee contracts, i.e. Sidama washed coffee and different origin washed coffee contracts respectively. R^2 results of experiment-2 and experiment-3.2 are relatively smaller.

Table 5.7: Analysis of MSE

S.N	Experiment	**MLP Model**				**RBF Model**			
		Minimum	Maximum	Average	Variance	Minimum	Maximum	Average	Variance
1	Experiment-1	0.0028	0.1047	0.0258	0.0271	0.0760	0.6199	0.1975	0.1477
2	Experiment-2	0.0578	0.1171	0.0794	0.0192	0.0681	0.3251	0.2016	0.0879
3	Experiment-3.1	0.0109	0.1402	0.0490	0.0467	0.0634	0.3753	0.1666	0.0882
4	Experiment-3.2	0.0288	0.1984	0.0673	0.0396	0.0807	0.9940	0.1935	0.2265

MSE Analysis by Method: - In all experiments, the smallest MSE with small variance is observed in MLP based models compared to models constructed by RBF algorithm (Table 5.7).

MSE Analysis by Experiment:- In MLP base models the smallest MSE with minimum variance is achieved Experiment 3.1. Second best result is observed in Experiment-1. Both coffee contracts are washed family. Relatively larger MSE and variance results were observed in the Experiment 2 and 3.2 both are of unwashed processing type.

From the analysis of results we have found that the selected algorithms MLP and RBF can model all export coffee contracts with varying degree of accuracy. In all experiments, MLP based models outperform RBF algorithm. This is due to the global approximation ability of MLP algorithm that means they are capable of generalizing even for little training data [23, 24]

The second finding of the study is prediction by coffee processing type. WSDA3 is of washed processing type; experiments 1 and 3.1 are also washed coffee contracts and the two experiments achieved higher performance result as compared to the others. From the result, it is concluded that the reference contract better predicts the washed processing coffee contracts than unwashed coffee contracts. This is due to the wide price variation between washed and unwashed coffee contract. Unwashed contracts are sold by a lesser price due to the additional cost required for unwashed contracts to change into washed coffee type after transaction is made. The researchers presumed that this price variation may be attributable to low prediction result.

CHAPTER SIX

CONCLUSION AND RECOMMENDATION

6.1. Conclusion

The goal of this research is to examine the price relationship between ECX coffee contracts and to propose a predictive model that represents the relationship.

The CRISP data model was used to undertake the research process. It has six phases: business understanding, data understanding, data preparation, model building, model evaluation and model deployment.

In the first phase, business understanding phase, the Ethiopia Commodity Exchange trading system was analyzed to articulate the problem.

In the second phase, five years and 6 months daily coffee trade data were collected and it is used for model building and testing the goodness of the models. To understand the general behavior of the data different descriptive statistics were computed.

To answer the first research question of the study "Is there a contract that can represent all other contracts" detailed relationship analysis between coffee contracts was undertaken by considering volume of trade, availability of contract parameters to determine the active contracts. About 12 contracts were identified and their relationship with the rest contract was investigated by descriptive statistics, scatter plot matrix and correlation analysis to select the reference contract. WDSA3 was found to be the best contract with positive relationship with almost all contracts.

The third phase is data preparation which is dedicated to make the collected data suitable for the model building process. In this phase, the major tasks that are undertaken are normalization and scaling. The daily closing price data are normalized into the range [-1, 1] while the textual warehouse information was mapped to the range [-1, 1].

The fourth phase is modeling, which includes the model selection, setting parameters for the selected model and the model building. ECX coffee contracts price data have non-linear pattern like most agricultural commodity prices. It is difficult to predict this data using the traditional methods. To overcome this problem, machine learning models are used as an alternative to complex and non-linear forecast models. Machine learning models include the artificial neural networks, support vector machines and relevance vector machines.

From the comprehensive literature survey artificial neural network is selected for this study because it can model flexibly linear or non-linear relationship among variables. Moreover, ANNs are also capable to approximate any continuous function to any desired accuracy. The two commonly used ANN prediction algorithms MLP and RBF neural network were used to model the price relationship between ECX coffee contracts. Finally, a three layered MLP neural network with two inputs neuron, one output neuron, and ten neurons in the hidden layer and with hyperbolic transfer functions were used. Similarly, RBF network with 2 inputs and one output neuron and spread value of 0.001 were used to build the models.

The fifth phase, is model evaluation, the performance of the constructed models were evaluated by R^2 and MSE. The model with larger R^2 and smaller MSE was selected.

From the analysis of results MLP and RBF algorithms have modeled all export coffee contracts with varying degree of accuracy. Comparing the empirical results of MLP with RBF neural network; the MLP neural network has minimum MSE and large R^2 values. This is because MLP network can generalize the general pattern in the data even for small sample size and RBF models are unable to make generalizations on small samples.

Generally, washed coffee contracts were predicted by the models constructed by MLP algorithm with higher accuracy while unwashed coffee contracts can also be predicted with an acceptable error range.

We have found that WSDA3 can be used as a reference contract to be traded in the market and the market will use this contract and the proposed MLP prediction model to determine the daily selling price of other export coffee contracts especially washed coffee contracts.

As a result:

- The market will have efficient trading system.
- WSDA3 price will be disseminated and it represents the real price
- It will be a good starting point for ECX futures market plan.

6.2. Recommendation and Future work

- Deploying the proposed MLP coffee contracts prediction model.

- The proposed model can also be used for unwashed contracts with minimum error. To improve the prediction performance of unwashed coffee contracts, we propose ULK6 (Selected as the second best contract in Section 3.1.7 and this contract has good correlation with all unwashed contracts, Annex 8) based prediction model need to be developed.

- To make ECX trading system more efficient; we recommend this study need to be carried out for other commodities such as sesame and haricot bean that are actively traded in the market.

- Globally, coffee market use ICO market daily price i.e. the Coffee C daily average price as a benchmark for their decision making. To provide Ethiopian coffee exporter intelligent market decision we recommend the relationship between ECX Export coffee contract and Coffee C daily price need to be analyzed and a predictive model that could guide the direction of trade should be developed in the future.

References

[1]. Adebiyi Ayodele A., Ayo Charles K., Adebiyi Marion O., and Otokiti Sunday O, *"Stock Price Prediction using Neural Network with Hybridized Market Indicators",* Journal of Emerging Trends in Computing and Information Sciences, VOL. 3, NO. 1, January 2012.

[2]. Aghababaeyan, Reza, and NajeebAhmadKhan TamannaSiddiqui. "*Forecasting the Tehran Stock Market by Artificial Neural Network."* International Journal of Advanced Computer Science and Applications Special Issue on Artificial Intelligence. 2011.

[3]. Berry, Michael, and Gordon Linoff. *Mastering data mining: The art and science of customer relationship management*. John Wiley & Sons, Inc., 1999.

[4]. Bishop, Chris. "*Improving the generalization properties of radial basis function neural networks."* Neural Computation 3, no. 4 (1991): 579-588.

[5]. Budhani, Neelima, and SANDEEP K. BUDHANI. *"Application of Neural Network in Analysis of Stock Market Prediction."* International Journal Of Computer science And Engineering Technology 3, no. 4 (2012).

[6]. Chapman, Pete, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, and Rudiger Wirth. *"CRISP-DM 1.0 Step-by-step data mining guide."* (2000).

[7]. Coffee Trade: New York Coffee Exchange. http://www.coffeeresearch.org/market/coffeemarket.htm.

[8]. Coffee Exporter's Guide. (2012). The International Trade Centre (ITC). http://www.intracen.org/The-Coffee-Exporters-Guide---Third-Edition.

[9]. Collard, J. E. "Commodity Trading with a Three Year Old. Neural Networks in Neural Networks in Finance and Investing - Using Artificial Intelligence to Improve Real-Word Performance." Probus Publishing Company, Chicago (1993).

[10]. Cybenko, George. *"Approximation by superpositions of a sigmoidal function."* Mathematics of control, signals and systems 2, no. 4 (1989): 303-314.

[11]. Dominic Stanculescu, Surendra Kotecha, Ethiopia Coffee Quality Improvement Project; www.intracen.org, Geneva, (2011).

[12]. Ebrahimi, Somayeh, Shahrokh Shajari, and Mohammad Hassan Tarazkar. *"Prediction of Agricultural Commodity Price Using Artificial Neural Networks: Case of Chicken Price in Fars province, Iran."* Journal of Basic and Applied Scientific Research 2, no. 11 (2012): 11537-11541.

[13]. Ethiopia Commodity Exchange. (2009), Company profile. http://www.ecx.com.et/CompanyProfile.aspx

[14]. Ethiopia Commodity Exchange Proclamation. Proclamation No. 550/2007. (2007); http://www.ecx.com.et/downloads/rules/ecexproclamation.pdf

[15]. Ethiopia Commodity Exchange Coffee Contract. http://www.ecx.com.et/downloads/Contracts/Coffee/CoffeeContracts.pdf

[16]. Fadlalla, Adam, and Chien-Hua Lin. *"An analysis of the applications of neural networks in finance."* Interfaces 31, no. 4 (2001): 112-122.

[17]. FAO, Grading and Classification of Green Coffee, http://www.fao.org/waicent/faoinfo/foodsafetyquality/cd_hygiene/cnt/cnt_en/sec_3/docs_3.3/Grading%20&%20class.pdf

[18]. Funahashi, Ken-Ichi. *"On the approximate realization of continuous mappings by neural networks."* Neural networks 2, no. 3 (1989): 183-192.

[19]. Gabre-Madhin, Eleni Z., and Ian Goggin. "*Does Ethiopia need a commodity exchange? An integrated approach to market development."* Addis Ababa: Ethiopian Development Research Institute, Working Paper Series 4 (2005).

[20]. Gabre-Madhin, Eleni. *" A market for Abdu: Creating a commodity exchange in Ethiopia. "* Intl Food Policy Res Inst, 2012.

[21]. Girma Nigussie, Kinato. *"Marketing information operation in Ethiopia with special reference to the Ethiopia Commodity Exchange (ECX) coffee trading."* (2011).

[22]. Han, Jiawei, and Micheline Kamber. Data Mining, Southeast Asia Edition: Concepts and Techniques. Morgan kaufmann, 2006.

[23]. Hornik, Kurt, Maxwell Stinchcombe, and Halbert White. *"Multilayer feedforward networks are universal approximators."* Neural networks 2, no. 5 (1989): 359-366.

[24]. Hornik, Kurt. *"Approximation capabilities of multilayer feedforward networks."* Neural networks 4, no. 2 (1991): 251-257.

[25]. Hornik, Kurt. *"Some new results on neural network approximation."* Neural Networks 6, no. 8 (1993): 1069-1072.

[26]. Hastie, Trevor, Robert Tibshirani, Jerome Friedman, T. Hastie, J. Friedman, and R. Tibshirani. The elements of statistical learning. Vol. 2, no. 1. New York: Springer, 2009.

[27]. ICE Coffee Brochure: Coffee "C" ..https://www.theice.com/publicdocs/ICE_Coffee_Brochure.pdf

[28]. Imam, Ibrahim F. *"An empirical study on the incompetence of attribute selection criteria."* In Foundations of Intelligent Systems, pp. 458-467. Springer Berlin Heidelberg, 1996.

[29]. Irie, Bunpei, and Sei Miyake. *"Capabilities of three-layered perceptrons." InNeural Networks*, 1988., IEEE International Conference on, pp. 641-648. IEEE, 1988.

[30]. Jackson, Joyce. *"Data Mining; A Conceptual Overview."* Communications of the Association for Information Systems 8, no. 1 (2002): 19.

[31]. Jain, Anil K., Jianchang Mao, and K. M. Mohiuddin. *"Artificial neural networks: A tutorial."* Computer 29, no. 3 (1996): 31-44.

[32]. Kara, Yakup, Melek Acar Boyacioglu, and Ömer Kaan Baykan. *"Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange."* Expert systems with Applications 38, no. 5 (2011): 5311-5319.

[33]. Kaastra, Iebeling, and Milton Boyd. *"Designing a neural network for forecasting financial and economic time series."* Neurocomputing 10, no. 3 (1996): 215-236.

[34]. Karlik, Bekir, and A. Vehbi Olgac. *"Performance analysis of various activation functions in generalized MLP architectures of neural networks."* International Journal of Artificial Intelligence and Expert Systems 1, no. 4 (2011): 111-122.

[35]. Kecman, Vojislav. Learning and soft computing: support vector machines, neural networks, and fuzzy logic models. MIT press, 2001.

[36]. Kohzadi, Nowrouz, Milton S. Boyd, Bahman Kermanshahi, and Iebeling Kaastra. *"A comparison of artificial neural network and time series models for forecasting commodity prices."* Neurocomputing 10, no. 2 (1996): 169-181.

[37]. Kolb, Robert W. *"Understanding Futures Markets."* Fifth Edition, Blackwell Publishers Ltd (1997)

[38]. Larose, Daniel T. Discovering knowledge in data: an introduction to data mining. John Wiley & Sons, 2014.

[39]. Lendasse, Amaury, John Lee, Eric De Bodt, Vincent Wertz, and Michel Verleysen. *"Approximation by radial basis function networks."* In Connectionist Approaches in Economics and Management Sciences, pp. 203-214. Springer US, 2003.

[40]. MathWorks. 2013. Functions in Neural Network Toolbox. http://www.mathworks.com/help/nnet/functionlist-alpha.html?searchHighlight=Functions%20in%20Neural%20Network%20Toolbox

[41]. McCulloch, Warren S., and Walter Pitts. *"A logical calculus of the ideas immanent in nervous activity."* The bulletin of mathematical biophysics 5, no. 4 (1943): 115-133.

[42]. Müezzinoğlu, Mehmet Kerem. *"Performance evaluation of recurrent RBF network in nearest neighbor classification."* In Artificial Intelligence and Neural Networks, pp. 58-69. Springer Berlin Heidelberg, 2006.

[43]. Naeini, Mahdi Pakdaman, Hamidreza Taremian, and Homa Baradaran Hashemi. *"Stock market value prediction using neural networks."* In Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference on, pp. 132-136. IEEE, 2010.

[44]. Neeraj, Mohan, Jha Pankaj, A. K. Laha, and Goutam Dutta. Artificial neural network models for forecasting stock price index in Bombay Stock Exchange. No. WP2005-10-01. Indian Institute of Management Ahmedabad, Research and Publication Department, 2005.

[45]. Park, Jooyoung, and Irwin W. Sandberg. *"Universal approximation using radial-basis-function networks."* Neural computation 3, no. 2 (1991): 246-257.

[46]. Mishra, Pragnyaban, Neelamadhab Padhy, and Rasmita Panigrahi. "The survey of data mining applications and feature scope." Asian Journal of Computer Science & Information Technology 2, no. 4 (2013).

[47]. Perwej, Yusuf, and Asif Perwej. *"Prediction of the Bombay Stock Exchange (BSE) market returns using artificial neural network and genetic algorithm."* (2012).

[48]. Priddy, Kevin L., and Paul E. Keller. Artificial neural networks: an introduction. Vol. 68. SPIE Press, 2005.

[49]. Rules of the Ethiopia Commodity Exchange; Rev. No. 5/2010. http://www.ecx.com.et/downloads/rules/ECXRuleAmharicEnglishFinal3.pdf

[50]. Shahwan, T., & Odening, M. (2007). Forecasting agricultural commodity prices using hybrid neural networks. In Computational intelligence in economics and finance (pp. 63-74). Springer Berlin Heidelberg.

[51]. Ticlavilca, Andres M., Dillon M. Feuz, and Mac McKee. *"Forecasting Agricultural Commodity Prices Using Multivariate Bayesian Machine Learning Regression Regression."* (2010).

[52]. Tong, Xiaojun, Zhu Wang, and Haining Yu. *"A research using hybrid RBF/Elman neural networks for intrusion detection system secure model."* Computer Physics Communications 180, no. 10 (2009): 1795-1801.

[53]. Witten, Ian H., and Eibe Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2005.

[54]. Yildiz, Birol, Abdullah Yalama, and Metin Coskun. *"Forecasting the Istanbul stock exchange national 100 index using an artificial neural network."* World Academy of Science, Engineering and Technology 46 (2008): 36-39.

[55]. Zhang, Guoqiang, B. Eddy Patuwo, and Michael Y Hu. *"Forecasting with artificial neural networks:: The state of the art."* International journal of forecasting 14, no. 1 (1998): 35-62.

[56]. Zou, H. F., G. P. Xia, F. T. Yang, and H. Y. Wang. *"An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting."* Neurocomputing 70, no. 16 (2007): 2913-2923.

Annexes

Annex 1: ECX Export coffee contracts

Annex 2: Coffee trade data total list of attributes

I.No	Attributes	Description
1	timedate	The Time the trade executes in seconds
2	tday	Trade date
3	tmonth	Trade month
4	tyear	Trade Year
5	DateKey	The unique identifier for the date variable
6	Product Type	The type of a product traded in the market coffee, sesame and haricot bean
7	Origin	The coffee origins, Jimma, limmu, yirgachefe and so on
8	ProcessingType	Washed and unwashed
9	MarketType	Export, local and speciality ypes
10	Name	Name of a class Washed Limmu, Unwashed Lekempti …
11	Class	Symbolic representation of class,
12	Symbol	Contract representation eg. WSDA6
13	Grade	Coffee grades 1 to 9 and undergrads
14	Openning price	The maket openning price
15	Closing price	The closing price of a day
16	Day high price	The highest price in a day
17	Day low price	The smallest price registed in a day
18	Diff	The difference between day high and day low
19	Volume in lot	Volume of coffee measured in lot
20	Volume in lot	Volume of coffee measured in quintal
21	Total value	Closing price multiplied by volume in lot
22	Production Year	The coffee production year
23	Warehouse	The place where coffee are stored

Annex 3: Descriptive summary statistics contracts

Annex 3.1: Washed and unwashed sidama coffee summary statistics

Contracts	Observation	Mean	Std. Dev.	Min	Max
usda3	51	1,036	364	420	1,560
usda4	116	963	368	345	1,610
usda5	324	948	385	312	1,640
usda6	541	962	363	301	1,620
usda7	356	881	357	300	1,580
usda8	73	759	321	320	1,450
usda9	20	701	325	305	1,335
usdb3	21	749	352	418	1,320
usdb4	55	497	175	300	1,200
usdb4	55	497	175	300	1,200
usdb5	83	538	272	310	1,400
usdb6	189	827	395	283	1,575
usdb7	144	832	309	282	1,460
usdb8	39	817	312	290	1,440
usdc3	94	1,153	303	385	1,560
usdc4	114	1,021	352	292	1,550
usdc5	79	791	356	335	1,540
usdc6	58	835	339	305	1,490
usdc7	34	913	281	330	1,365
usdc8	12	746	332	250	1,250
usdcug	26	445	127	275	900
usdd5	55	851	329	455	1,455
usdd6	72	1,043	327	480	1,585
usdd7	26	1,235	270	810	1,580
wsda3	680	967	313	500	1,600
wsda3	680	967	313	500	1,600
wsda4	552	921	314	487	1,585
wsda5	249	872	305	462	1,500
wsda6	182	805	291	465	1,450
wsda7	66	745	271	450	1,290
wsda8	27	683	211	445	1,180
wsdaug	70	425	65	330	540
wsdaugnp	62	717	165	460	1,240
wsdb3	600	906	300	406	1,555
wsdb4	577	905	309	426	1,600
wsdb5	227	864	285	455	1,450
wsdb6	125	799	279	460	1,470
wsdb6	125	799	279	460	1,470
wsdb7	64	792	281	470	1,400

Contracts	Observation	Mean	Std. Dev.	Min	Max
wsdb8	27	706	221	435	1,170
wsdb9	25	687	284	368	1,100
wsdbug	72	436	68	323	550
wsdbugnp	54	727	174	455	1,210
wsdbugp	19	768	239	370	1,150
wsdc3	200	948	303	380	1,465
wsdc3	200	948	303	380	1,465
wsdc4	199	928	319	430	1,500
wsdc5	83	974	296	430	1,410

Annex 3.2: Descriptive statistics for washed coffee contracts of different origins.

Origin	Contracts	No. of Observation	Mean	Average price	Minimum price	Maximum price
Bebeka	wbb3	113	930.24	236.76	610.00	1,370.00
	wbb4	86	953.01	224.50	580.00	1,410.00
Lekemte	wlk3	127	935.87	262.48	570.00	1,520.00
	wlk4	110	953.61	273.92	548.00	1,460.00
Limmu	wlma3	277	1,029.20	198.72	548.00	1,460.00
	wlma4	218	1,004.45	201.99	497.00	1,320.00
	wlma5	93	965.95	214.27	600.00	1,300.00
	wlmb4	93	928.72	261.78	550.00	1,500.00
	wlmb5	45	952.94	234.71	580.00	1,500.00
Tepi	wtp3	105	920.74	211.90	495.00	1,276.00
Yirgachefe	wyca3	364	1,197.83	217.48	694.00	1,830.00
	wyca4	133	1,158.04	227.13	740.00	1,765.00
	wycb3	336	1,192.48	208.90	610.00	1,750.00
	wycb4	226	1,146.70	220.82	600.00	1,700.00
	wycb5	78	1,041.71	264.71	580.00	1,570.00
	wycb6	55	1,002.56	232.53	655.00	1,360.00

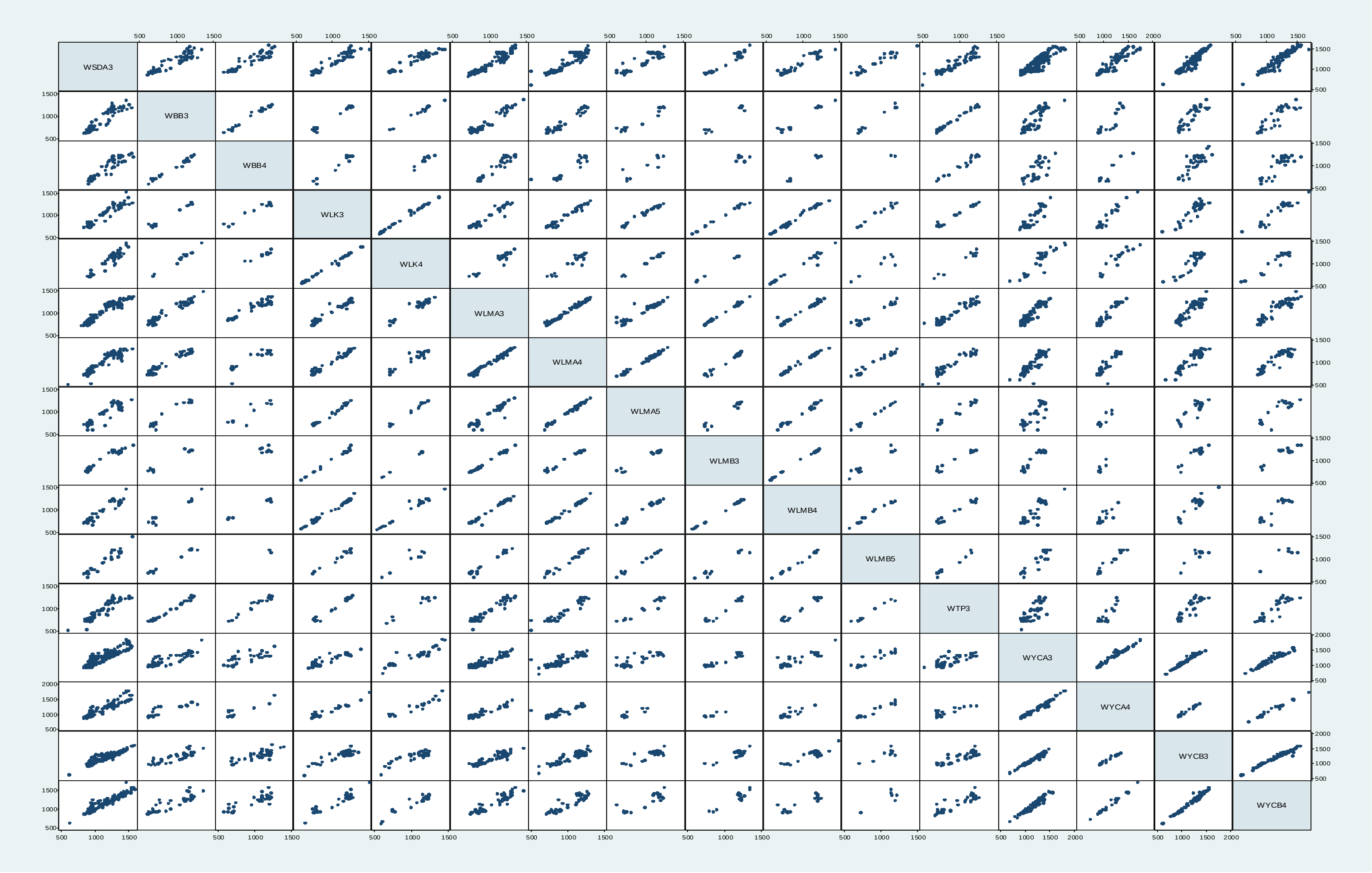
Annex 3.3: Descriptive statistics for unwashed of different origin

Origin	Contracts	No. of Observation	Mean	Average price	Minimum price	Maximum price
Forest	ufra6	114	880.51	230.41	410.00	1,380.00
	ufra7	340	897.59	229.44	385.00	1,475.00
	ufra8	610	873.90	231.27	395.00	1,470.00
	ufrb8	219	950.77	222.41	450.00	1,450.00
	ufrb7	106	999.81	213.79	720.00	1,555.00
	ufrb9	132	935.92	219.89	650.00	1,350.00
Jimma	ujma6	472	985.83	195.04	730.00	1,520.00
	ujma7	672	979.11	208.81	542.00	1,535.00
	ujma8	628	965.34	204.39	670.00	1,505.00
	ujma9	280	1,034.18	213.68	430.00	1,465.00
	ujmb6	748	950.49	246.20	410.00	1,578.00
	ujmb7	778	944.42	250.14	389.00	1,555.00
	ujmb8	547	914.03	263.23	390.00	1,510.00
	ukf6	120	925.75	134.22	720.00	1,315.00
	ukf7	169	888.53	147.41	700.00	1,300.00
	ukf7	169	888.53	147.41	700.00	1,300.00
Lekemte	ulk3	238	920.03	286.23	490.00	1,580.00
	ulk4	517	956.63	269.56	475.00	1,580.00
	ulk5	707	961.38	263.96	465.00	1,585.00
	ulk6	821	951.45	258.12	450.00	1,580.00
	ulk7	794	943.77	253.68	440.00	1,560.00
	ulk8	538	947.98	245.95	430.00	1,555.00
	ulk9	215	948.78	254.36	400.00	1,520.00

Annex 4: Correlation WSDA3 with washed other Origins coffee

	WSDA3	*WLK3*	*WLK4*	*WLK5*	*WLMA3*	*WLMA4*	*WLMA5*	*WLMA6*	*WLMB4*	*WLMB5*	*WLMB6*	*WTP3*	*WTP4*	*WYCA3*	*WYCA4*	*WYCA5*	*WYCB3*	*WYCB4*
WSDA3	1.000																	
WLK3	0.953	1.000																
WLK4	0.925	0.996	1.000															
WLK5	0.932	0.992	0.999	1.000														
WLK6	1.000	1.000																
WLMA3	0.957	0.976	0.960	0.965	1.000													
WLMA4	0.946	0.985	0.922	1.000	0.996	1.000												
WLMA5	0.939	0.991	0.994		0.963	0.995	1.000											
WLMA6	0.924	0.891	0.954	1.000	0.924	0.934	0.984	1.000										
WLMB4	0.952	0.994	0.986	0.996	0.989	0.997	0.992	0.880	1.000									
WLMB5	0.951	0.985	0.910		0.985	0.985	0.995	0.994	0.993	1.000								
WLMB6	0.953	0.993	1.000		0.936	0.987	0.927		1.000	1.000	1.000							
WTP3	0.935	0.989	0.977	1.000	0.964	0.950	0.969	0.867	0.984	0.981		1.000						
WTP4	0.883	0.947	0.997		0.903	0.902	(0.013)		0.904	0.786		0.951	1.000					
WYCA3	0.892	0.926	0.913	0.991	0.936	0.902	0.871	0.798	0.881	0.893	0.984	0.792	0.573	1.000				
WYCA4	0.935	0.960	0.927		0.950	0.890	0.866	0.912	0.837	0.971	0.840	0.838	(0.938)	0.983	1.000			
WYCA5	0.824		0.999		1.000	1.000								0.999	0.997	1.000		
WYCB3	0.948	0.908	0.878	0.679	0.922	0.925	0.882	0.868	0.899	0.846	1.000	0.839	0.709	0.988	0.988		1.000	
WYCB4	0.953	0.944	0.953	0.962	0.933	0.918	0.887	0.928	0.889	0.843		0.876	0.873	0.962	0.991	0.914	0.988	1.000
WYCB5	0.890	0.978	0.933	0.381	0.967	0.950	0.959	0.896	0.981	0.979	1.000	0.974	1.000	0.932	0.984		0.910	0.976
WYCB6	0.930	0.968	0.955		0.917	0.977	0.982	0.933	0.883			0.979	1.000	0.908	0.970		0.899	0.978

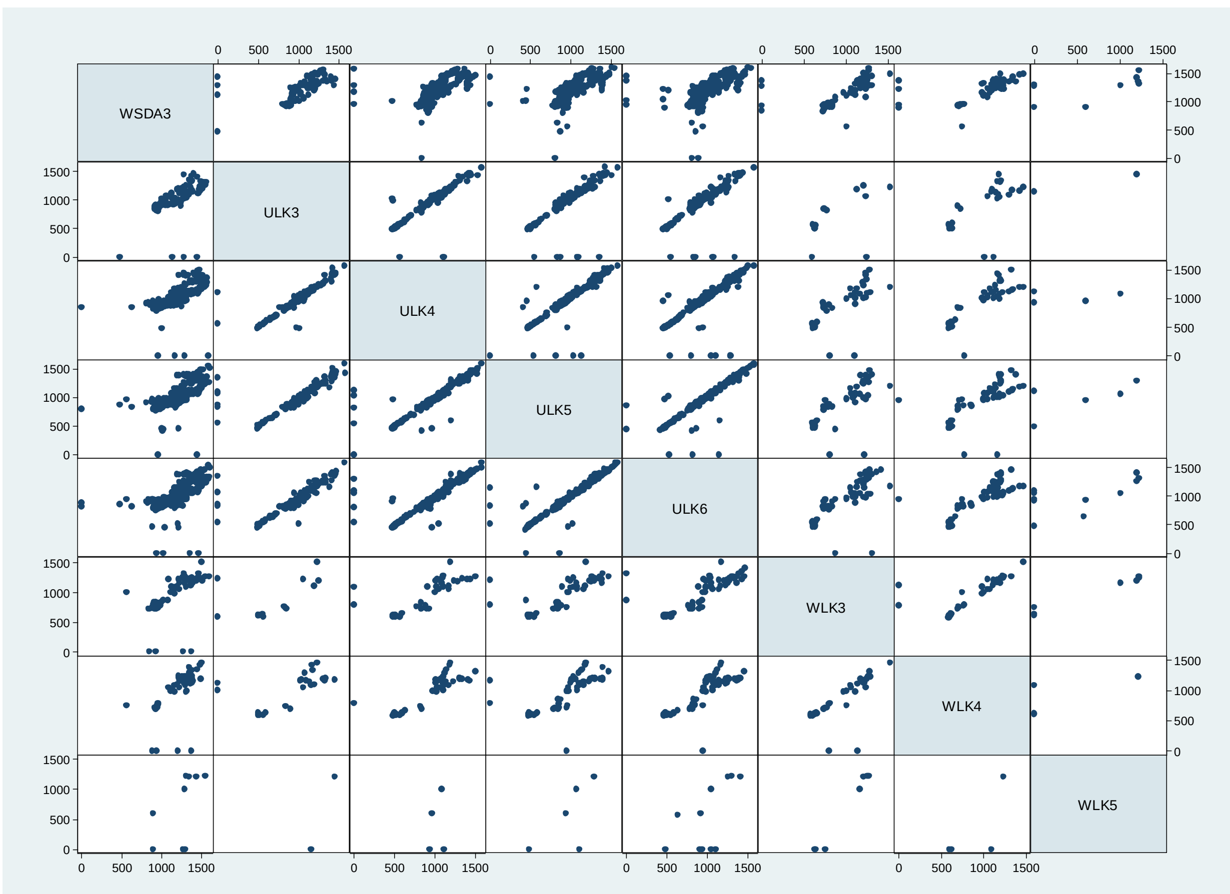
Annex 5: Correlation WSDA3 with washed other Origins coffee



Annex 6: Correlation WSDA3 with unwashed coffee origins

	wsda3	ufra7	ufra8	ufra9	ufraug	ujma6	ujma7	ujma8	ujma9	ujmb4	ujmb5	ujmb6	ujmb7
wsda3	1												
ufra7	0.7589	1											
ufra8	0.8216	0.9741	1										
ufra9	0.8539	0.9136	0.9582	1									
ufraug	0.8725	0.9058	0.9595	0.9904	1								
ujma6	0.7413	0.8705	0.8912	0.8413	0.8191	1							
ujma7	0.7718	0.9911	0.9731	0.9121	0.8981	0.91	1						
ujma8	0.4809	0.7244	0.6934	0.691	0.6519	0.6421	0.7536	1					
ujma9	0.6174	0.6516	0.6567	0.5901	0.5697	0.5207	0.6515	0.45	1				
ujmb4	0.2835	0.0493	0.1404	0.3523	0.3356	0.0672	0.0286	0.0136	0.0147	1			
ujmb5	0.8443	0.947	0.9593	0.9421	0.9388	0.8563	0.9581	0.7249	0.6454	0.0476	1		
ujmb6	-0.0232	-0.1295	-0.2223	-0.3145	-0.3337	-0.0011	-0.096	-0.1651	-0.221	-0.3969	-0.1995	1	
ujmb7	0.7948	0.988	0.9794	0.9369	0.9223	0.8888	0.9937	0.7513	0.6921	0.0665	0.9762	-0.1701	1

Annex 7: Scatter plots WSDA3 with unwashed coffee of different origins



Annex 8: ULK6 correlation with unwashed coffee contracts

	UFRA5	**UFRA6**	**UFRA7**	**UFRA8**	**UFRB6**	**UFRB7**	**UFRB8**	**UFRB9**	**UJMA6**	**UJMA7**	**UJMA8**	**UJMA9**	**UJMB6**	**UJMB7**	**UJMB8**	**UKF6**	**UKF7**	**ULK3**	**ULK**
UFRA5	1.000																		
UFRA6	1.000	1.000																	
UFRA7	0.997	0.998	1.000																
UFRA8	0.990	0.991	0.993	1.000															
UFRB6	0.968	0.997	0.984	0.978	1.000														
UFRB7	0.998	0.991	0.989	0.986	1.000	1.000													
UFRB8	0.985	0.990	0.995	0.994	0.990	0.997	1.000												
UFRB9	0.994	0.990	0.993	0.994	0.997	0.988	0.997	1.000											
UJMA6	0.990	0.990	0.989	0.980	0.979	0.985	0.986	0.986	1.000										
UJMA7	0.989	0.987	0.988	0.981	0.986	0.991	0.987	0.986	0.997	1.000									
UJMA8	0.992	0.990	0.992	0.989	0.988	0.990	0.994	0.994	0.990	0.991	1.000								
UJMA9	0.990	0.988	0.989	0.989	0.985	0.989	0.990	0.993	0.980	0.979	0.994	1.000							
UJMB6	0.996	0.992	0.989	0.983	0.989	0.991	0.986	0.986	0.996	0.995	0.987	0.975	1.000						
UJMB7	0.997	0.991	0.991	0.986	0.992	0.993	0.988	0.988	0.996	0.996	0.990	0.980	0.999	1.000					
UJMB8	0.998	0.992	0.993	0.991	0.991	0.986	0.993	0.991	0.993	0.994	0.995	0.990	0.996	0.997	1.000				
UKF6	0.986	0.985	0.989	0.965	0.897	0.951	0.929	0.910	0.984	0.983	0.984	0.960	0.986	0.987	0.985	1.000			
UKF7	0.992	0.994	0.995	0.978	0.843	0.995	0.995	0.986	0.989	0.988	0.989	0.981	0.988	0.989	0.986	0.997	1.000		
ULK3	0.995	0.989	0.983	0.977	0.947	0.993	0.979	0.972	0.980	0.980	0.969	0.957	0.994	0.993	0.993	0.948	0.967	1.000	
ULK4	0.994	0.984	0.983	0.972	0.998	0.990	0.981	0.976	0.983	0.983	0.970	0.964	0.992	0.992	0.992	0.960	0.962	0.999	1.000
ULK5	0.992	0.986	0.985	0.976	0.996	0.990	0.983	0.974	0.986	0.986	0.977	0.968	0.994	0.994	0.993	0.965	0.972	0.997	0.999
ULK6	0.994	0.987	0.985	0.977	0.994	0.992	0.985	0.976	0.986	0.986	0.979	0.971	0.994	0.994	0.993	0.971	0.973	0.997	0.998
ULK7	0.992	0.987	0.984	0.977	0.992	0.992	0.986	0.977	0.989	0.987	0.982	0.979	0.994	0.994	0.993	0.969	0.973	0.997	0.997
ULK8	0.994	0.989	0.986	0.984	0.988	0.991	0.990	0.983	0.986	0.988	0.986	0.980	0.991	0.993	0.994	0.926	0.955	0.990	0.993
ULK9	0.998	0.977	0.988	0.988	0.988	0.992	0.991	0.987	0.985	0.988	0.989	0.985	0.986	0.989	0.993	0.825	0.914	0.990	0.986
USDA4	0.982	0.849	0.822	0.848	0.113	0.916	0.953	0.957	0.860	0.892	0.872	0.865	0.879	0.882	0.903	0.900	0.853	0.940	0.930

	UFRA5	UFRA6	UFRA7	UFRA8	UFRB6	UFRB7	UFRB8	UFRB9	UJMA6	UJMA7	UJMA8	UJMA9	UJMB6	UJMB7	UJMB8	UKF6	UKF7	ULK3	ULK
USDA5	0.932	0.904	0.929	0.916	0.998	0.969	0.962	0.978	0.924	0.942	0.944	0.941	0.941	0.940	0.947	0.851	0.903	0.955	0.955
USDA6	0.944	0.923	0.945	0.946	0.989	0.969	0.976	0.977	0.951	0.958	0.962	0.951	0.961	0.960	0.970	0.902	0.929	0.965	0.955
USDA7	0.990	0.975	0.968	0.959	0.956	0.956	0.981	0.986	0.962	0.964	0.970	0.958	0.965	0.970	0.975	0.841	0.935	0.964	0.969
USDB4	(1.000)		0.972	0.768			0.986	1.000	0.743	0.810	0.977	1.000	0.883	0.851	0.772	0.975	0.985	0.984	0.874
USDB5	1.000	0.990	0.966	0.952		0.992	0.804	0.872	0.916	0.930	0.921	0.826	0.946	0.938	0.944	0.829	0.948	0.969	0.965
USDB6	0.967	0.953	0.950	0.953	0.911	0.948	0.978	0.975	0.931	0.943	0.946	0.923	0.948	0.948	0.966	0.945	0.946	0.979	0.968
USDC5		0.919	0.951	0.917	1.000	0.887	0.894	0.941	0.875	0.856	0.904	0.865	0.851	0.842	0.813	0.789	1.000	0.817	0.792
UYCA5	0.994	0.970	0.969	0.940	(1.000)	0.958	0.977	0.990	0.912	0.931	0.947	0.941	0.915	0.927	0.949	0.734	0.935	0.985	0.985
UYCA6	1.000	0.962	0.959	0.948	1.000	0.992	0.986	0.986	0.952	0.937	0.963	0.973	0.956	0.946	0.975	0.972	0.978	0.975	0.957
UYCB6	0.942	0.940	0.949	0.944	0.994	0.957	0.965	0.985	0.929	0.943	0.950	0.932	0.944	0.946	0.962	0.893	0.920	0.976	0.969

Annex 9: ECX coffee warehouse scaled representation

I. N	Coffee Contract	Warehouse name code	Warehouse name code
1	WSDA3	Addis Ababa	[1]
2	WSDA4, WSDA5, WSDA6, WSDA7, WSDBYGNP	Awassa, Addis Ababa	[-1,1]
3	WSDB3, WSDB4	Awassa, Dilla	[-1,1]
4	WSDB5, WSDB6, WSDB7 WSDBYGNP	Awassa	[1]
5	WSDC3 , WSDC4, WSDC5, WSDC6	Welayta Sode	[1]

Annex 10: WDSA4 prediction result on test data set and Denormalized price data

Transformed Data			Renormalized to price		
target	**Predicted BY MLP**	**Predicted by RBF**	**TEST TARGET**	**Predicted BY MLP**	**Predicted by RBF**
(0.544)	(0.576)	(0.577)	943.00	930.32	930.00
(0.402)	(0.444)	0.021	1,000.02	983.41	1,170.04
0.121	0.138	0.102	1,210.00	1,217.04	1,202.50
0.096	0.048	0.096	1,200.00	1,180.67	1,200.00
0.270	0.237	0.258	1,269.99	1,256.72	1,265.00
0.245	0.226	0.202	1,259.99	1,252.16	1,242.50
0.347	0.454	0.026	1,300.98	1,343.86	1,172.04
0.395	0.467	0.021	1,320.01	1,349.12	1,170.04
0.470	0.489	0.451	1,350.00	1,357.84	1,342.50
0.594	0.643	0.021	1,399.99	1,419.80	1,170.04
0.345	0.283	0.245	1,300.02	1,275.22	1,260.00
0.208	0.283	0.245	1,245.01	1,275.22	1,260.00
0.034	(0.006)	(0.019)	1,174.99	1,159.08	1,154.00
0.308	0.267	0.263	1,285.00	1,268.50	1,267.20
0.395	0.392	0.402	1,320.01	1,318.84	1,323.00
0.395	0.341	0.021	1,320.01	1,298.54	1,170.04
0.445	0.440	0.395	1,340.01	1,338.04	1,320.00
0.499	0.483	0.021	1,362.01	1,355.47	1,170.04
0.519	0.479	0.507	1,370.00	1,353.77	1,365.00
0.228	0.261	0.021	1,253.00	1,266.34	1,170.04
0.208	0.248	0.244	1,245.01	1,260.94	1,259.43
0.245	0.288	0.024	1,259.99	1,277.16	1,171.14
0.245	0.267	0.263	1,259.99	1,268.50	1,267.20
0.245	0.267	0.263	1,259.99	1,268.50	1,267.20
0.258	0.267	0.263	1,265.01	1,268.50	1,267.20
0.245	0.283	0.245	1,259.99	1,275.22	1,260.00
(0.228)	(0.039)	(0.004)	1,070.00	1,145.79	1,160.00
(0.527)	(0.458)	(0.444)	949.99	977.71	983.33
(0.079)	(0.084)	(0.029)	1,129.98	1,127.92	1,150.00
(0.352)	(0.424)	(0.471)	1,020.01	991.11	972.50
(0.377)	(0.347)	(0.284)	1,010.01	1,022.10	1,047.50
(0.602)	(0.653)	(0.651)	920.00	899.45	900.00
(0.589)	(0.661)	(0.662)	925.02	896.30	895.82
(0.539)	(0.622)	(0.626)	945.01	911.62	910.00
(0.527)	(0.606)	(0.670)	949.99	918.32	892.50
(0.499)	(0.554)	(0.549)	960.99	938.97	941.00
(0.502)	(0.505)	(0.514)	959.99	958.54	955.00
(0.726)	(0.683)	(0.651)	870.01	887.44	900.00
(0.626)	(0.642)	(0.624)	910.00	903.93	910.78
(0.616)	(0.626)	(0.689)	914.02	910.15	885.00
(0.739)	(0.683)	(0.651)	864.99	887.44	900.00
(0.739)	(0.661)	(0.662)	864.99	896.30	895.82
(0.626)	(0.626)	(0.689)	910.00	910.15	885.00
(0.602)	(0.606)	(0.670)	920.00	918.32	892.50

(0.589)	(0.581)	(0.558)	925.02	928.20	937.50
(0.564)	(0.581)	(0.558)	935.01	928.20	937.50
0.096	0.226	0.202	1,200.00	1,252.16	1,242.50
(0.402)	(0.375)	(0.386)	1,000.02	1,010.93	1,006.67
(0.664)	(0.689)	0.014	894.98	884.72	1,167.10
(0.801)	(0.816)	0.021	840.02	833.91	1,170.04
(0.651)	(0.663)	0.015	900.00	895.31	1,167.36
(0.714)	(0.727)	(0.720)	874.99	869.65	872.50
(0.714)	(0.727)	(0.720)	874.99	869.65	872.50
0.046	(0.006)	(0.019)	1,180.01	1,159.08	1,154.00
0.726	0.731	0.021	1,452.99	1,455.06	1,170.04
0.943	0.912	0.021	1,539.99	1,527.72	1,170.04
0.046	(0.006)	(0.019)	1,180.01	1,159.08	1,154.00

Annex 11: ULK7 prediction result on test data set and Denormalized price data

Actual	Predicted BY MLP	Predicted by RBF	Actual TEST Data	Predicted BY MLP	Predicted by RBF
(0.4865)	(0.4735)	(0.0856)	940.00	944.81	1,088.33
(0.5135)	(0.5266)	(0.0856)	930.00	925.17	1,088.33
(0.4730)	(0.6213)	(0.0856)	945.00	890.11	1,088.33
(0.5811)	(0.6942)	(0.6622)	905.00	863.15	875.00
(0.5541)	(0.7605)	(0.5541)	915.00	838.60	915.00
(0.6811)	(0.6155)	(0.7230)	868.00	892.27	852.50
(0.6081)	(0.7376)	(0.7198)	895.00	847.10	853.67
(0.6622)	(0.8180)	(0.8108)	875.00	817.35	820.00
(0.7297)	(0.6036)	(0.4507)	850.00	896.66	953.25
(0.6622)	(0.3834)	(0.1301)	875.00	978.14	1,071.86
(0.5811)	(0.5509)	(0.0856)	905.00	916.16	1,088.33
(0.3378)	(0.1484)	(0.2419)	995.00	1,065.10	1,030.50
0.1351	0.0380	0.2696	1,170.00	1,134.05	1,219.75
0.1081	0.2274	0.0405	1,160.00	1,204.13	1,135.00
-	0.3616	0.3784	1,120.00	1,253.78	1,260.00
0.1351	0.4059	0.4865	1,170.00	1,270.20	1,300.00
0.2162	0.3725	0.8716	1,200.00	1,257.82	1,442.50
0.3784	0.3675	0.4054	1,260.00	1,255.96	1,270.00
0.3378	0.2274	0.0405	1,245.00	1,204.13	1,135.00
0.5811	0.3572	0.3122	1,335.00	1,252.18	1,235.50
0.5946	0.3725	0.8716	1,340.00	1,257.82	1,442.50
0.4459	0.3272	0.2014	1,285.00	1,241.06	1,194.50
0.3784	0.3672	(0.0856)	1,260.00	1,255.87	1,088.33
(0.1486)	(0.1224)	(0.0703)	1,065.00	1,074.73	1,094.00
(0.3297)	(0.5036)	(0.0915)	998.00	933.67	1,086.15
(0.5405)	(0.1739)	(0.2888)	920.00	1,055.66	1,013.14
(0.9189)	(0.5020)	(0.5865)	780.00	934.27	903.00
(0.6622)	(0.3108)	(0.4932)	875.00	1,004.99	937.50
(0.7568)	(0.6511)	(0.5743)	840.00	879.08	907.50
(0.7432)	(0.7216)	(0.7838)	845.00	853.02	830.00
(0.2432)	(0.0154)	(0.0709)	1,030.00	1,114.31	1,093.75
(0.5135)	(0.5301)	(0.6149)	930.00	923.86	892.50
(0.5676)	(0.5800)	(0.0856)	910.00	905.40	1,088.33

Annex 12: Creating Neural Networks

Annex 12.1: MLP neural network

➢ **Network creation**

This step is started with selection of network architecture. In this research the first ANN model we are going to use is the MLP network, its code is:-

net = newff(PR,[S1 S2...SNl],{TF1 TF2...TFNl},BTF,BLF,PF)

- PR - Rx2 matrix of min and max values for R input elements.
- Si - Size of ith layer, for Nl layers.
- TFi - Transfer function of ith layer, default = 'tansig'.
    - BTF - Backprop network training function, default = 'trainlm'.
    - BLF - Backprop weight/bias learning function, default = 'learngdm'.
    - PF - Performance function, default = 'mse'
- newff : create and returns "net" = a feed-forward backpropagation network.

➢ **Network Initialization**

Network initialization initialises the networks weight and biases. It is one means of improving the performance of the network by retraining the built network each time after retraining the network there is need to initialize the original network.

```
>> net = init(net);      % init is called after newff
```

The overall architecture of your neural network is store in the variable net; variable can be reset

net.trainParam.epochs =1000;(Max no. of epochs to train) [100]

net.trainParam.lr =0.001;   (learning rate, not default trainlm) [0.01]

- net.trainFcn=trainlm ; a variant of BP based on second order algorithm (*Levenberg-Marquardt)*
- TRAIN trains a network NET according to NET.trainFcn and NET.trainParam

TRAIN(NET,P,T,Pi,Ai)

- NET - Network.
- P - Network inputs.
- T - Network targets, default = zeros.
- Pi - Initial input delay conditions, default = zeros.
- Ai - Initial layer delay conditions, default = zeros.

➢ Simulation of the network

[Y] = SIM(model, UT)

- Y : Returned output in matrix or structure format.
- model : Name of a block diagram model.
- UT : For table inputs, the input to the model is interpolated.

➢ Evaluate Goodness of Fit (Performance Evaluation)

- Comparison between target and network's output in testing set.
- Comparison between target and network's output in training set.
- Measure the distance/similarity of the target and output, or simply use *mse.*

Annex 12.2: Code for creating MLP neural network

net= *newrb*(*P, T, GOAL,SPREAD, MN, DF*) ……// network creation

Where *P* is input sample,

T is expected response,*G*

GOAL is training precision,

SPREAD is a density of RBF and its default value is 1,

MN is the maximum number of neurons,

DF is the increasing number of neurons (Geand Sun, 2007).

Among these, *SPREAD* is an important parameter. Set to 0.001.

Declaration

This thesis is my original work and has not been submitted as a partial requirement for a degree in any university.

____

Frehiwot Mulugeta Deressa

The thesis has been submitted for examination with my approval as university advisor.

____

Dr. Sebisbe Hailemariam

Addis Ababa, Ethiopia

October 2014