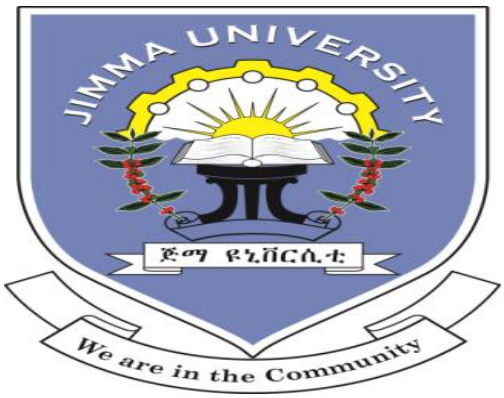


JIMMA UNIVERSITY

COLLEG OF NATURAL SCIENCES

DEPARTMENT OF INFORMATION SCIENCE

PREDICTING MATERNAL MORTALITY RATE USING DATA MINING TECHNIQUES: THE CASE OF JIMMA UNIVERSITY SPECIALIZED HOSPITAL MATERNITY WARDS.

BY:

EDOSA FEKADU (BSC).

Principal Advisor: Dr Million Meshesha (Ph.D).

Co-Advisor: Ms. Elsabet Wedajo (Msc).

June, 2017
Jimma, Ethiopia

JIMMA UNIVERSITY

COLLEG OF NATURAL SCIENCES

DEPARTMENT OF INFORMATION SCIENCE

PREDICTING MATERNAL MORTALITY RATE USING DATA MINING TECHNIQUES: THE CASE OF JIMMA UNIVERSITY SPECIALIZED HOSPITAL MATERNITY WARDS.

BY

EDOSA FEKADU

A Thesis Submitted to the Department of Information Science of Jimma University in partial fulfilments for the Degree of Master of Science in Information Science

Principal Advisor: Dr Million Meshesha (Ph.D)

Co-Advisor: Ms. Elsa Wedajo (Msc.)

June, 2017

Jimma, Ethiopia

JIMMA UNIVERSITY

COLLEG OF NATURAL SCIENCES

DEPARTMENT OF INFORMATION SCIENCE

Predicting Maternal Mortality Rate Using Data Mining Techniques: The case of Jimma University Specialized Hospital Maternity wards.

BY

EDOSA FEKADU

Members of the examining board:

Name	Title	Signature	Date
_____	Chair person	_____	_____
_____	Co-Advisor	_____	_____
_____	Advisor	_____	_____
_____	Internal-Examiner	_____	_____
_____	External-Examiner	_____	_____

DECLARATION

I declare that the thesis is my original work and it has not been presented for a degree in any other university. All the material sources used in this work are duly acknowledged.

_____

EDOSA FEKADU

June, 2017

This thesis has been submitted to the department for examination with our approval as University advisors:

Principal Advisor: Dr Million Meshesha (Ph.D) _____

Co-Advisor: Elsabet Wodajo (MSC.) _____

June, 2017

DEDICATION

I dedicate this Thesis to Almighty God my creator, my strong pillar, my source of inspiration, wisdom, knowledge and understanding. He has been the source of my strength throughout this program and on his wings only have I soared. I also dedicate this work to my Biological Mam Ayantu which I lost her when I was grade twelve; Helen Ing who had encouraged me all the way and whose encouragement has made sure that I give it all it takes to finish that which I have started. To Monica Barlow, who is helping me and encouraging me when I was in University. Lastly, all my friends who supported me during this paper. God bless you!!

ACKNOWLEDGMENT

Frist and Foremost, I would like to express my sincere gratitude to my advisor Dr. Million Meshesha for the continuous support of my research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my study.

Next to my advisor MS Elsabet Wedajo for her guidance and hard working with me, As well as I would like to thank the rest of my department which gives this wonderful chance for me. Secondly people who work in Jimma University specialized hospital maternity wards:

To Dr Dajane Tolasa, for his encouragement, insightful comments and hard questions. My sincere thanks also goes to Ing Family who supports me during my study in Information and Knowledge Management. Beside this support, love, commitment and advising from Ing family.

Last but not the least; I would like to thank my mam even though she is not alive: my mother Ayantu, for giving birth to me at the first place and supporting me spiritually throughout my life until I grow up.

Contents page

LIST OF ACCRONYMS AND ABBREVIATIONS

ANN	Artificial Neural Network
ARFF	Attribute relation File Format
CSV	Comma separated Format
DHS	Demographic and Health Surveys
GBD	Global Burden of Disease
JUSHMW	Jimma University Specialized Hospital Maternity ward
KDD	Knowledge Discovery in Data Base
KNN	K-Nearest Neighbor
MOTHERS BP	Mothers Blood Pressure
MWW	Maternity World Wide
NGO	Non-Governmental Organization
PART	Partial Decision Tree
SEMMA	Sample Explore Modify Model Asses
SPSS	Statistical Package for Social Science
UNFPA	United Nations Population Fund
WEKA	Waikato Environment for Knowledge Learning
WHO	World Health Organization

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Maternal mortality is the death of women during pregnancy, childbirth or in the 42 days after delivery remains a major challenge to health systems as a worldwide. Recent reports from WHO and UNAIDS indicate that the number pregnant woman died after delivery is increasing from time to time. This number is dramatically increasing in sub Saharan African countries including Ethiopia.

The main objective of this study is to develop a predictive model for the maternal mortality status. The overall activity of this study is guided by a Hybrid-DM processes model and used the data of Jimma University specialized Hospital maternity ward. The study has used 4218 instances, seven predicting and one outcome variables to run the experiments. The mining algorithms; J48 decision tree, Naïve Bayes and PART rule induction are used in all experiments due to their popularity in recent related works. Ten-fold cross validation and 70/30 split criteria test option were used to train and test the classifier models. J48 decision tree algorithms were better performance with 98.74 % accuracy running on 10 fold cross validation test option with default parameter using 14 attribute than any experimentation done in this study. The selected attributes is significant for maternal mortality rate status which is after delivery has been identified. Those are Mothers BP, Address, APGAR score, Diagnosis, Mothers Age, Length of stay, Indication and Condition on Discharge.

A promising result is observed in applying data mining techniques to build predictive model for maternal mortality using socio-demographic, clinical and biological features. This study is proved that the prediction of maternal mortality can be applicable with help of data mining application in the maternity ward data and predicting the life status of the mothers after delivery had been identified. This study did not include pregnant women life expectancy, Therefore, for the future work developing a model which could predict the life expectancy of pregnant women after labor and delivery would need further study beside the developed model as well as the improvement of J48 models.

Keywords: Maternal mortality, Hybrid DM process, Predictive Model, Classification, PART Rule Induction, J48 Decision tree, Naïve Bayes.

CHAPTER ONE

1.1 BACKGROUND

Data Mining is becoming popular in healthcare field because there is a need of efficient analytical methodology for detecting unknown and valuable information in health data (Ranjani, 2013). In health industry, Data Mining provides several benefits such as detection of the fraud in health insurance, availability of medical solution to the patients at lower cost, detection of causes for diseases and identification of medical treatment methods. There is a need to generate a powerful tool for analyzing and extracting important information from this complex data. The analysis of health data improves the healthcare by enhancing the performance of patient management tasks (Agarwal, 2013).

Health care industry today generates large amounts of complex data about patients, hospital resources, disease diagnosis, electronic patient records, medical devices, drugs and human resources. Hospital is the one among the health station that give the service of health. Larger amounts of data are a key resource to be processed and analyzed for knowledge extraction that enables support for cost-savings and decision making that can be achieved through different types of techniques (Ranjani, 2013).

Data mining can be applied for investigation of problems that occur in health sector. The Healthcare industry is among the most information intensive industries. Medical information, knowledge and data keep growing on a daily basis. It has been estimated that an acute care hospital may generate five terabytes of data a year. The ability to use these data to extract useful information for quality healthcare is crucial. Maternal mortality can be applicable by the new modern techniques of data mining. Data mining is one among the most important steps in the knowledge discovery process. It can be considered the heart of the knowledge discovery process (Dharani , 2014).

Women's are more vulnerable for death than men's due to different types of health related problems. This might be due to lack of treatment in the health station or level of life. In developing country the life status and cultural related staff can be a reason of women's death too. From this the researcher could understand that maternal death is the big percentage in the developing country. The Jimma university specialized hospital is one of the health organization

that provide a health service for the community. Jimma university Hospital has different types of wards which gives services for the patient depending on the cases of the patient. There are wards in the hospital related with treatment cases like female ward, male ward, pediatric ward, isolation room and maternity ward. Maternity wards are wards that give service of labor and deliver for women's. This ward also financed with maternity worldwide with corporation of Jimma University. The main aim of this ward is to reduce deaths of mothers and help the risk that comes through labor and deliver. Pregnancy women needs special care and the new baby should be physically and mentally health. Maternal mortality is the death of women during pregnancy, childbirth, or in the 42 days after delivery remains a major challenge to health systems worldwide. Global initiatives to intensify policy intervention for maternal mortality began with the Safe Motherhood Initiative in (1987) (Margaret.et al., 2010).

Pregnant womens are helped from the maternity ward by getting service freely for the purpose of labor and delivery.The maternity ward is an organization which works to save the life of pregnant women during labor and delivery. As well as minimizing the factors wich could affect mothers life after they addimitted from this ward makes still necessary to save the life of womans. Therefore, predicting the maternal mortality rate status after they came to the ward would make womans to live in safety and confidencial life status.

1.2 OVERVIEW OF JIMMA UNIVESRITY SPECIALISED HOSPITAL

Jimma University Hospital is one of the health institution found in Jimma town which is emerged with the emergence of Jimma University institution. Known by the Jimma University hospital before advanced to specialization Hospital. The Jimma University Specialized Hospital is working for the community by giving the health service.

The Jimma University specialized hospital gives the service of saving life for the whole community not only people who lives in Jimma. Also Jimma University Specialized Hospital works with different types of NGO to support the life safe task. Maternity ward is one of the NGO that works with Jimma University specialized hospital to safe the life of pregnant women. As a general the mission and vision of Jimma University is to work for the people of Ethiopia and providing health care service.

1.3. STATEMENT OF THE PROBLEM

WHO indicates that maternal mortality rate has decreased as worldwide. However, maternal mortality remains high in low income countries in which 99% of the deaths occur. Of this, Sub-Saharan Africa alone shoulders three fifth of maternal deaths (Gedefaw et al, 2014). Reliable information about rates and trends in maternal mortality is essential for resource mobilization, for planning and assessment of progress towards. In view of the continued prominence of maternal mortality as a health and development goal, global rates and trends in maternal mortality need to be reassessed. Several efforts have been made over nearly three decades to improve the quality of information about maternal mortality, including the incorporation of sibling history modules in the Demographic and Health Surveys (DHS) and similar surveys the inclusion of questions about whether recent deaths were related to pregnancy in censuses and the use of record linkage or confidential enquiry to identify under-registration of maternal deaths in vital registration systems. According to the report from WHO, reducing maternal mortality is not just an issue of development, also an issue of human rights too and as preventable maternal mortality often represents a violation of woman's right to life.

Women's are dying from severe bleeding, infections, eclampsia, obstructed labor and from consequences of unsafe abortions and a majority of all these causes are preventable. Maternal mortality is the leading cause of death for 15-19 years and adolescent girls in developing countries. This indicates that, in cross-national regressions, having trained attendants at delivery is strongly associated with lower maternal mortality levels.Women's are dying needlessly and suffering disabling conditions as a result of pregnancy and childbirth (Shiffman , 2000).

Jimma University specialized hospital has wide range of wards which is categorized depending up on to the services. It has been stated that managing all wards information and gain one valuable information is lacked. Health Information Management today generates large amounts of complex data about patients, hospital resources, disease diagnosis, electronic patient records and medical devices. Large amount of data are a key resource to be processed and analyzed for knowledge extraction that enables support for cost-savings and decision making (Durairaj & Ranjani, 2013).

Each section of Jimma University hospitals ward has a record. There is potentially useful hidden knowledge in those records. Maternity ward is one of the important wards that Jimma university specialized hospital works with. The maternity worldwide and the hospital need a clear information or knowledge about the issue of maternal mortality rate. Since Jimma University specialized hospital is wide and information rich there is no valid knowledge in maternity ward. The information that is gathered in case of pregnancy women in the ward is stored but there is no analyzed information, filtered knowledge, non-patterned knowledge and non-identified attribute wich could predict pregnant woman. The stored data is huge in this ward but it is difficult to see the pattern. Since the data is not classified or clustered the probability of using the hidden knowledge is very low.

Encoding the data at every single time of labor and deliver which is done by maternity ward workers. This record is stored after the pregnant women got a service from the maternity ward. But there is no counseling depending on the information which is stored for the pregnant women even though by applying data mining would be possible. It is clear that to hospitalize one pregnancy women during her first visit, the ward registers her information. Analyzing the mortality rate of mothers and the case status of her future life is the crucial problem for the Jimma University Specialized Hospital maternity wards. The reason why maternal rate is increase is so complex for the chief executor organization (CEO) of the hospital to know means weather the case is because of infrastructure, professional, transportation, attention, service and Lack of counseling.

Predicting the futurity of mother's mortality rate and provide decision was the issue for the Jimma University Maternity ward. Because of this the Ethiopian health minister could not handle and solve the case of maternal problem which is really important for the developing country like Ethiopia. There is plenty of record in the ward but it is a big puzzle for the Jimma university specialized hospital to identify and improve the reasons that pregnant women die period and report in case for the higher authority. Jimma University specialized hospital could not get hidden and potentially useful knowledge from this ward to improve and help maternal cases in case of minimizing and improving the life of pregnant women after pregnant woman served.

Different researchers tried to apply data mining for prediction of mortality on maternal. Tesfahun, (2012) tried to predict adult mortality by using data mining, the developed model does not included maternal mortality rate and depends on some socio demographic data too. The developed model does not include maternal mortality rate which means does not include women's during labor and delivery. This could be as a post factors determinant and pre factors determinant. It only works for both sex and the age of the adult should be also more than thirty. But with this maternal mortality woman's got pregnancy after fifteen years in average. The case of their death during pregnancy in the ward is not determined and which factors would enforce women's to death during labor and delivery is still the left for further study.

Dawit, (2013) also shows that predicting maternal health care seeking pattern of reproductive age, that level of education and the number of children that she has as the attributes for the model and her pre life style. It does not concern the condition of woman's in delivery and labor in wards. There is no identified factor that could predict about the death of woman's during labor and deliver in the maternity ward and after they go back to their house, rather it concerns factors for their life expectancy and mortality as a general instead of developing a model for pregnant woman. Maternal mortality death is very hot at the age of 15-19 during labor and delivery in the ward, not only above 15 because there is a chance that could women's got pregnancy below 15 years old. That is to mean the more they are young the more they are forced to death.

According to Gedefaw.et al, 2014 data on maternal near miss cases and events among mothers who received care at health institutions is lacking and maternal mortality is high but still this lacks what types of attributes could affect mother life at health care. There is no identified attributes and which part of health sector could affect mother's life after they got service. Beside it is not applied by data mining application. The gap from the above literature review the predicting maternal mortality rate for female starting from fifteen years old and which attribute would determine the life of pregnant women in the future or attributes which predicts their death is un touched part during labor and delivery.

Therefore, the aim of this study is to apply data mining for constructing a predictive model for determining maternal mortality rate. To this end, this study attempts to explore and answer the following research question

- What are the relevant attributes for predicting a maternal mortality using data mining techniques?
- Which data mining algorithm is the good predictor to solve the problem of maternal mortality rate?

1.4. OBJECTIVE OF THE STUDY

1.4.1. GENERAL OBJECTIVE

The general objective of this research is to develop a predictive model for the maternal mortality by using data mining techniques.

1.4. 2.SPECIFIC OBJECTIVES

Even though the big aim of this research is building a predicting model for maternal mortality, there are specific objectives formulated to achieve the main objective.

- ✓ To categorize attributes which could affect the maternal mortality.
- ✓ To prepare quality datasets for training and testing
- ✓ To demonstrate which algorithm of the classification algorithm is best to predict maternal mortality rate.

1.5. SIGNIFICANCE OF THE STUDY

The study provides supports for the stated problems about the life status of pregnant women after she served from the wards. The significance of the maternal mortality rate prediction is the major one in health environment to minimize the death of pregnant women after she admitted from the health care. This helps the countries minister of health, policy makers, workers, health professionals and the people of the country for the sustainability life of women's. The significance of the study can be multiple but it also provides clear and fluent form of information which generated from the model. This study also significant to provide a physician to counseling the pregnant women's after labor and delivery by using the developed system on

the visual basic interface. Helps the physicians of maternity wards by providing the knowledge of counseling after the pregnant woman got service. Beside this, providing fluent report for higher authority, Jimma University specialized hospital CEO and Creating awareness for the people of Ethiopia and Jimma community, additionally it includes the regions, woreda's and kebele consequently. It is also good for guiding the health professional and hospitals. It helps the non-governmental organization to provide more attention on the maternal cases. As well as, good for NGO which like to improve care for mothers. Generally, after the completion the study, the model will provide a support on the maternal case for health worker and government to understand the mortality rate and it helps the whole hospital of Jimma university.

1.6. SCOPE AND LIMITATIONS OF THE STUDY

The scope of this research is limited to predict maternal mortality rate using the data that found in Jimma University specialized hospital "maternity ward database". The model which only predicts maternal mortality rate without including other wards of Jimma University specialized hospital. The limitation of this study is the scarcity of encoded data while accessing the saved data from the maternity ward.

1.7. RESEARCH METHODOLOGY

The overall activity of this thesis were used a Hybrid model which is six step knowledge discovery process model. The Hybrid DM process model describes procedures that are performed in each of its steps (Jinhong et al., 2009). Due to the nature of the problem and attributes in the dataset, classification mining task were selected to build the classifier models. J_{48} decision tree, Naïve Bayesian and PART rule induction algorithm is used for this study for developed the model. It is undoubted that the data mining techniques has different types of models which can solve the real world problem which is really integrated with the data base, record, data ware house and files. The explanations of those models are discussed under literature review and chapter three of this study. The Hybrid model contains both the KDD and CRISP-DM at one place. Therefore, the researcher used the Hybrid DM processes model for this specified study. Because, this model is really fruit full in the prediction of health oriented problem and has many option of feedback mechanism more than CRISP DM (Fayyad et al., 1998). The hybrid DM processes model provides the six step processes to generate the

important knowledge and model development. Therefore, the study is completed under the guidance of the Hybrid DM processes model.

1.7.1. RESEARCH DESIGN

This research is designed to identify the determinant of the maternal mortality rate. To explore the application of data mining on this particular research, hybrid (Ciso.et al) data mining methodology was employed. The explanations of those models are discussed under literature review and chapter three of this study. The hybrid model provides the six step processes to generate the important knowledge and model development i.e understanding of domain, understanding of data, preparation of data, data mining, evaluations of discovered knowledge and uses of knowledge.

1.7.2. DOMAIN UNDERSTANDING

This initial step involves working closely with domain experts to define the problem and determine the project goals, identifying key people, and learning about current solutions to the problem. It also involves learning domain-specific terminology. For this purpose the researcher is selected Doctors, mid-wife and Nurses as a domain expert by clarifying the problem and aim of the study.

Justifying the selected attribute for the stated problem regarding to the prediction of the maternal mortality rate is very essential. Factors or conditions which affect the life of the pregnant woman before and after the deliver is discussed by journals of WHO. HIV/AIDS and cardiac diseases are account for about one fifth of maternal deaths. According to the WHO, hypertension, which together account for more than half of maternal deaths Indirect causes, which include deaths, due to conditions such as malaria, Regional estimates show that hemorrhage and hypertension are the top listed risky condition for the maternal during pregnancy (WHO). The researcher listed out the above idea as HIV as NR and Cardiac diseases, hypertension as a mother's blood pressure (BP). Beside this the content of platelets or Trombocytopenia and the distance that pregnant ladies come also has an effect on the maternal deaths depending on the domain expert's recommendation. The attributes or the factors which predict the deaths of maternal mortality is additionally approved by domain experts too.

1.7.3. DATA UNDERSTANDING

This step includes collecting sample data and deciding which data, including format and size, will be needed. For this purpose Secondary data was used as a source of information. The secondary data included pregnant woman's data which is found in Jimma University Specialized Hospital maternity ward and the data is saved in the form of Excel.

Collecting representative subset of pregnant women data was a prerequisite to address the objective of this research. Therefore, data were collected from the records of pregnant women's who were in the Jimma university specialized hospital maternity ward called maternity ward data base. The data base was in Excel form through which the data clerks can enter pregnant woman data and generate different reports. A full backup of the database of the pregnant woman was taken from Jimma university specialized Hospital maternity ward. It stores data for pregnant woman's data during labor and delivery. The data set stored for three years record from the year 2014 to 2017. The total data set obtained on Maternity ward data base is around 5648 instances which is saved up to the end of april fifteen of 2017.

1.7.4. DATA PREPARATION

This step concerns deciding which data will be used as input for DM methods in the subsequent step. It involves sampling, running correlation and significance tests, and data cleaning, which includes checking the completeness of data records, removing or correcting for noise and missing values. This step covers preprocessing the data and making ready the maternity ward data for model development. For this study WEKA version 3.7.5 is the major tool to drive the study forward. WEKA is a good tool for the hybrid model since the model has different types of steps to preprocess the data. There is also other tool of data mining like Rapid miner and TANAGRA, but due to the researcher's familiarity with WEKA tool and WEKA tool has so many java dependent environment and preprocessing option more than Rapid miner and Tanagra. Also so many researchers have been used this tool before, for the purpose of mining hidden knowledge in the stored data. Additionally, MS-Excel is the primary tool which the used for the purpose of data cleaning and preprocessing situation.

1.7.5. DATA MINING

Here the data miner uses various DM methods to derive knowledge from preprocessed data. The selected algorithm for this study is J48 decision, PART rule induction and Naïve Bayes applied on the data to dig out the hidden knowledge by using Weka tool for the prediction maternal mortality rate.

1.7.8. EVALUATION

This is the fifth step of the Hybrid DM processes model which evaluates the discovered knowledge and rules are evaluated for the predicting maternal mortality. The performance and accuracy of the model on the data is investigated depending on the model responds and discussing with Domain expert weather the knowledge is valid for prediction maternal mortality.

1.7.8. USE OF KNOWLEDGE

This is the final step which the knowledge founded is implemented for the prediction of maternal mortality. This could be through developing a prototype for the stated problem by using different tools. Lastly the researcher used the Microsoft visual to develop the interface which could predict the maternal mortality rate of pregnant women. Submitting the final results and systems for the hospital would be the end.

1.8. ETHICAL CONSIDERATION

The ethical approval and clearance were obtained from Jimma University CNS and Postgraduate Programme Office. The necessary explanation about the purpose of the study and about its procedure, assurance of confidentiality, the right not to participate on the study without any consequences was included.

1.9. ORGANIZATION OF THESIS

This thesis is organized into six chapters. The first chapter deals with the general overview of the study including background, statement of the problem, research objectives, significance of the research, methodology, scope and limitations of the research. The second chapter focuses on literature review on maternal mortality and data mining concepts and techniques of this study. Literature reviews and related works by data mining on predicting maternal mortality. Chapter three is data mining Architecture and methods. Chapter four is about data preparation which constitutes business understanding, data understanding and data preprocessing. Therefore, at this stage of modeling a quality data is made for the classification algorithms. Chapter 5 is where the experiments conducted were presented. Here, topics about data mining, model selection and Prototype development were discussed in detail. Results of the experiments are also analyzed and interpreted and discussion of findings too is presented. Chapter 6 is the final chapter which presents concluding remarks and recommendations of the study.

CHAPTER TWO

LITERATURE REVIEW

2. 1.MATERNAL MORTALITY

Maternal mortality is defined as the death of a woman while pregnant or within 42 days of termination of pregnancy, regardless of the site or duration of pregnancy, from any cause related to or aggravated by the pregnancy or its management (WHO, 2015). Female literacy rates area strong Predictor of Maternal Mortality Rates. The more literate a female population, the lower the Maternal Mortality Rate appears. A woman's chance of dying or becoming disabled during pregnancy and childbirth is closely connected to her social and economic status, the norms and values of her culture, and the geographic remoteness of her home. That means if women live in the comfortable environment the chance of their death is very rare. Generally speaking, the poorer and more marginalized a woman is, the greater her risk of death. In fact, maternal mortality rates reflect disparities between wealthy and poor countries more than any other measure of health. A woman's lifetime risk of dying as a result of pregnancy or childbirth is 1 in 39 in Sub-Saharan Africa, as compared to 1 in 4,700 in industrialized countries. Globally, at least 585, 000 women die each year by complications of pregnancy and child birth More than 70% of all maternal deaths are due to five major complications: hemorrhage, infection, unsafe abortion, hypertensive disorders of pregnancy, and obstructed labor. The majority of maternal deaths (61%) occur in the postpartum period, and more than half of these take place within a day of delivery. Worldwide an estimated 500,000 women die as a result of pregnancy each year (Admasu, 2014).

2.2. MORTALITY RATE

Mortality rate or death rate is a measure of the number of deaths (in general, or due to a specific cause) in a particular population, scaled to the size of that population, per unit of time. Mortality rate is typically expressed in units of deaths per 1,000 individuals per year; thus, a mortality rate of 9.5 (out of 1,000) in a population of 1,000 would mean 9.5 deaths per year in that entire population, or 0.95% out of the total. It is distinct from "morbidity", a term used to refer to either the prevalence or incidence of a disease, and also from the incidence rate (the number of newly appearing cases of the disease per unit of time) (Noelle et al., 2013) .

2.3. MATERNAL MORTALITY AND CAUSES OF DEATH

Maternal mortality is a serious problem for under developing country. Maternal mortality is the leading in Africa and Asia. Since the launching of the Safe Motherhood Initiative in1987, there has been a worldwide effort to reduce maternal mortality and to identify its determinants. The declaration of the Millennium Development Goals (MDGs) aiming at reducing by three-quarters the maternal mortality ratio between 1990 and 2015 has also increased the demand for measuring maternal mortality at national and subnational levels (Ana .et al, 2005).Maternal mortality is notoriously difficult to measure..Every year, more than 289,000 women die during pregnancy or childbirth. Most of these deaths are preventable.At least 12 million women suffer severe maternal complications.The chance of dying is much greater in poor countries; developing countries account for 99 percent of the global maternal deaths, the majority of which are in sub-Saharan Africa and southern Asia (Abdella, 2010).

2.4. CAUSES OF MATERNAL MORTALITY IN ETHIOPIA

There are so many cuases for maternal mortality in ethiopia. As an example listing some points will be psossible like awwerness about health,early marriage,remotnes from health station,lack of food,transport and cost of service for pre checking during the pregnency. Globally, at least 585, 000 women die each year bycomplications of pregnancy and child birth. More than 70% of all maternal deaths are due to five major complications:hemorrhage, infection, unsafe abortion, hypertensive disorders of pregnancy and obstructed labor. The majority of maternal deaths (61%) occur in the postpartum period and more than half of these take place within a day of delivery (Ahmed , 2010).

The number of maternal deaths is highest in countries where women are least likely to have skilled attendance at delivery, such as a midwife, doctor or other trained health professional. Likewise, within countries, it is the poorest and least educated women who are most vulnerable to maternal death and disability.High maternal mortality rates are an indication not only of poorly functioning health systems, but also of deep-seated gender inequalities that leave women with limited control over decision-making and that restrict their access to social support, economic opportunities and health care.These gender inequalities manifest early in life; girls born into poverty are more vulnerable to child marriage and exploitation, such as sex trafficking or forced labor. Adolescent girls frequently lack the power to decide whether contraception is

used during sex, or whether sex takes places at all. This places them at high risk for early pregnancy and its resulting complications (UNFPA, 2012).

According to Ahmed (2010), Lack of information and adequate knowledge about danger signals during pregnancy and labor, cultural traditional practices that restrict women from seeking health care, lack of money, Out of reach of health facilities, poor road,communication network,community support Mechanisms Delay, Inadequate skilled attendants, poorly motivated staff, inadequate equipment and supplies weak referral system, procedural guides and the like are the causes for the mortality of maternal.

2.5. MATERNITY WARD AND THE GOALS

As the name implies maternity ward a ward which cares womens during pregnancy only. Labor and deliver is accomplished only in this ward. This ward is funded by a organization called maternity world wide (MWW). The maternity world wide is known by supporting womens with pregnancy. As well as this organization largely expanded in Africa and some Asian country. The main goal of this organization is to save womens life during pregnancy and support womens who could not afford pay during labor and delivery. Also to produce a generation for the future world which highly affected by poornes. A woman's relationship with maternity care providers and the maternity care system during pregnancy and childbirth is vitally important (WHO, 2010).

2.6. OVERVIEW OF DATA MINING CONCEPT

Data mining has attracted a great deal of attention in the information industry and insociety as a whole in recent years, due to the wide availability of huge amounts of data and the imminent need for turning such data into useful information and knowledge. Data mining can be viewed as a result of the natural evolution of information technology. Data mining requires identification of a problem, along with collection of data that can lead to better understanding, and computer models to provide statistical or other means of analysis. Data mining tools need to be versatile, scalable, capable of accurately predicting responses between actions and results, and capable of automatic implementation. Versatile refers to the ability of the tool to apply a wide variety of models. Scalable tools imply that if the tools works on a small data set, it should also work on larger data sets (Kamber, 2006).

Data mining is a powerful tool to dig out the hidden knowledge from a large repository of data. In section how data mining is applied in the health environment specifically to digging out hidden knowledge which is helpful for the purpose of predicting. Today data mining helps as a tool and a technique when our real world is becoming information rich in data and poor in information. Maurizio (2011) defined data mining as knowledge mining from data. And the other writers define data mining as knowledge extraction, information discovery, information harvesting, data archeology, and data pattern processing. Knowledge Discovery in Databases (KDD) is generally used to refer to the overall process of discovering useful knowledge from data and data mining is a particular step in this process. The term data mining can be also a synonymous with data dredging or data snooping and has been used to describe the process of trawling through data in the hope of identifying patterns with the statistician's perspective (Jackson, 2002).

Even though the name of the data mining is different, the task and function of data mining is still the same. The main aim of data mining is extracting useful knowledge from a large data base and data ware house. This means there is nothing that can limit data mining in direction in case of finding any hidden knowledge from any data which means in health sector, business and scientific purpose. Healthcare data mainly contains all the information regarding patients as well as the parties involved in healthcare industries. Due to continuous increasing the size of electronic healthcare data a type of complexity is exist in it. Data Mining mainly extracts the meaningful patterns which were previously not known. These patterns can be then integrated into the knowledge and with the help of this knowledge essential decisions can become possible (Parvez et al., 2015).

Data mining architecture containing many elements in it like Data Mining Engineer, Pattern evaluation, Data Warehouse server, User Interface and Knowledge Base. Data mining system of Architecture is not far from KDD procesess (Chaudhary, 2015).

Therefore, it is clear that data mining digs out data according to the importance the needed knowledge. This means data mining can be applicable with the techniques and methods of knowledge discovering within any stored data. The data mining is applicable in any direction and the architecture of the data mining looks like the figure below.

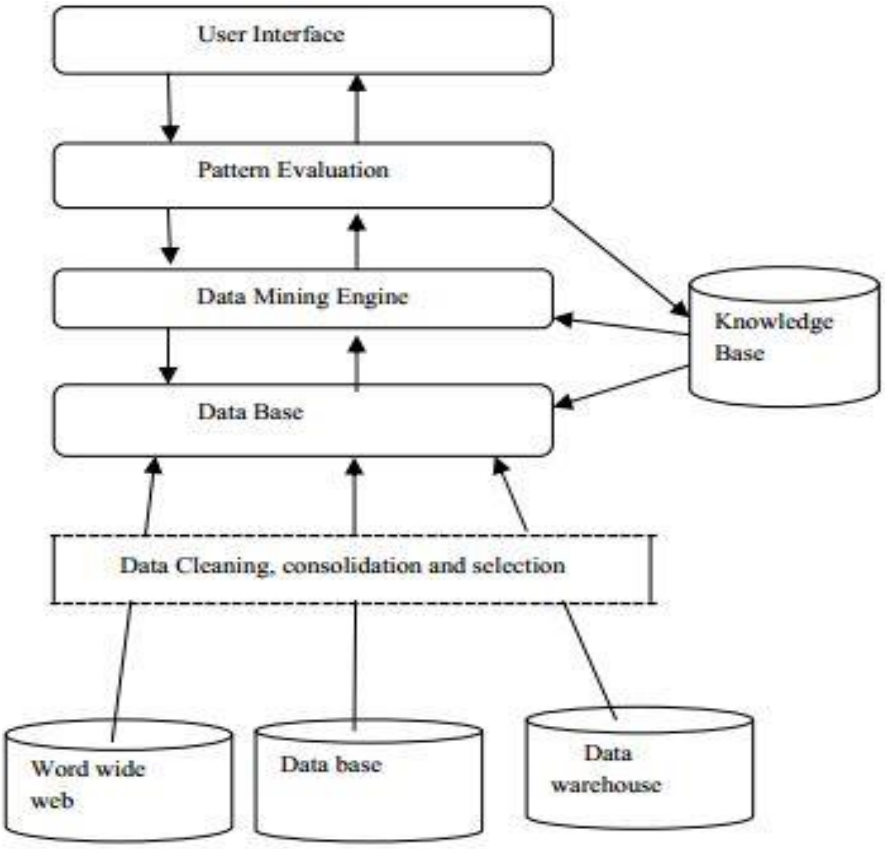


Figure 2.1: Data mining system architecture (Chaudhary, 2015).

2.7. DATA MINING PROCESSES

Data Mining is a process of discovering various models, summaries, and derived values from a given collection of data**.** Data mining process consists of three major steps. It all starts with a big pile of data. The first processing step is data preparation often referred to as "scrubbing the data". Data is selected, cleaned, and preprocessed under the guidance and knowledge of a domain expert. The most time-consuming part of the data mining process is preparing data for data mining. This step can be streamlined in part if the data is already in a database, data warehouse, or digital library, although mining data across different databases. Second, a data mining algorithm is used to process the prepared data, compressing and transforming it to make it easy to identify any latent valuable nuggets of information. In the second step in data mining,once the data is collected and preprocessed, the data mining algorithms perform the actual sifting process. Many techniques have been used to perform the common data mining

activities of associations, clustering, classification,modeling, sequential patterns, and time series forecasting (Sivanandam & Sumathi, 2006).

The third phase is the data analysis phase, where the data mining output is evaluated to see if additional domain knowledge was discovered and to determine the relative importance of the facts generated by the mining algorithms.The final step is the analysis of the data mining results or output. In some cases the output is in a form that makes it very easy to discern the valuable nuggets of information from the trivial or uninteresting facts. The relationships are represented in if-then rules form. With rules recast into textual form, the valuable information is much easier to identify. In other cases, however, the results will have to be analyzed either visually or through another level of tools to classify the nuggets according to the predicted value. Whatever be the data mining algorithm used, the results will have to be presented to the user. A successful data mining application involves the transformation of raw data into a form that is more compact and more understandable, and where relationships are explicitly defined (Sivanandam & Sumathi, 2006).

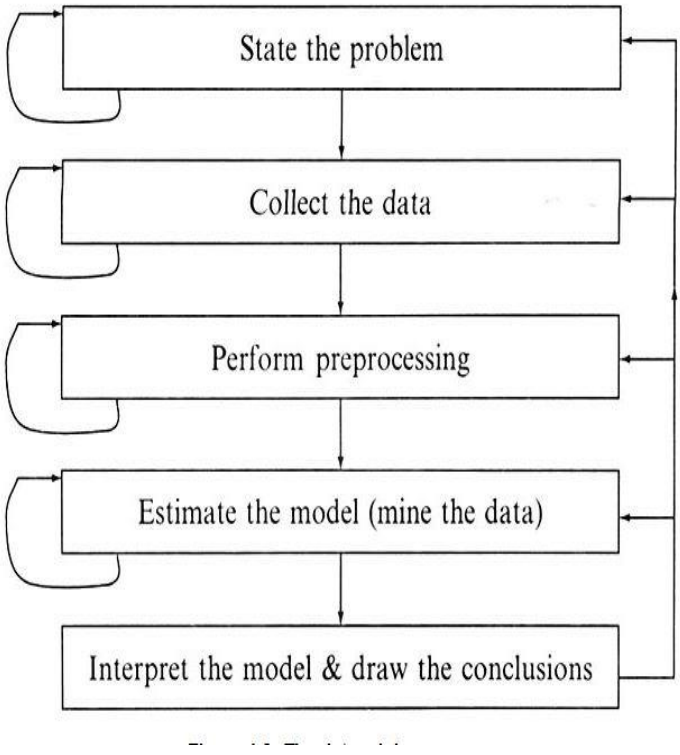


Figure 2.2: Data mining procesess (Sivanandam & Sumathi, 2006).

Figure 2.2 shows the all procesess of data mining procesess step by step and the data mining process iterates through five basic steps .

Data selection: This step consists of choosing the goal and the tools of the data mining process, identifying the data to be mined, then choosing appropriate input attributes and output information to represent the task.

Data transformation: Transformation operations include organizing data in desired ways, converting one type of data to another (e.g from symbolic to numerical), defining new attributes, reducing the dimensionality of the data, removing noise, outliers, normalizing, if appropriate, deciding strategies for handling missing data.

Data mining step-perse: The transformed data is subsequently mined, using one or more techniques to extract patterns of interest. The user can significantly aid the data mining method by correctly performing the proceeding steps.

Result interpretation and validation: For understanding the meaning of the synthesized knowledge and its range of validity, the data mining application tests its robustness, using established estimation methods and unseen data from the database. The extracted information is also assessed (more subjectively) by comparing it with prior expertise in the application domain.

Incorporation of the discovered knowledge: This consists of presenting the results to the decision maker who may check/resolve potential conflicts with previously believed or extracted knowledge and apply the new discovered patterns (Sivanandam & Sumathi, 2006).

2.8. DATA MINING PROCESSES MODEL

In order to systematically conduct data mining analysis, a general process is usually followed accordignly.This process provides a good coverage of the steps needed starting with data exploration, data collection, data processing, analysis, inferences drawn, and implementation. There are some standard processes CRISP-DM, Hybrid Model, KDD and SEMMA (Delen, 2008).

2.8.1. KNOWLEDGE DISCOVERY IN DATABASE (KDD)

The first KDD process was proposed by Fayyad in 1996. This process consists of several steps that can be executed iteratively. KDD has been more formally defined as it is non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data. KDD is the process of knowledge discovery while data mining is a technique applied or knowledge discovery considered as just a step in the entire process (Ana, 2008). As shown in Figure 2.2 the KDD process consists of five steps: data selection, data preprocessing, data transformation, data mining and interpretation/evaluation.

Given data, the first step in KDD is data selection. In this stage creating a target dataset on focus of a subset of variables needed on which discovery aimed to solve the problem are selected. For discovery purposes, data relevant to the analysis task are retrieved from the database and unnecessary data attributes should be removed. In order to produce effective data mining models in terms of quality and performance, the raw data need to undergo preprocessing in the form of data cleaning. Because real world data are mostly dirty and unclean which need to correct bad data that encountered from data redundancy, incompleteness or missing attributes value, noise, and inconsistency in order to make knowledge searching paths ease for mining algorithms. Therefore, data quality needs to be assured in this step before ahead to next phase of knowledge discovery process in data mining. Because of the use of different sources, data that is fine on its own may become problematic when we want to integrate it. In this step data need to be combined from multiple sources, such as database, data warehouse, files and non-electronic sources into a coherent store. We need to merge different sourced data by keeping uniform format for all before running data mining tools and techniques. During transformation phase, data are consolidated into forms appropriate for mining to reduce data size by dividing the range of data attribute into intervals each containing approximately same number of samples or to scale attribute data to fall within a specified range. Therefore, values of attributes are changed to a new set of replacement values to ease data mining. Data mining is the next essential process where intelligent methods are applied in order to extract hidden patterns in the data. This phase requires analysis of the main problem for patterns of interest in the data depending on the business objectives and data mining requirements. Different data mining algorithms and techniques are used for searching knowledge or interesting patterns to construct predictive or descriptive models (Ana, 2008).

Model creation is followed by performance evaluation which measures the accuracy rate of the system. The mined pattern enables to identify the truly interesting ones. For any errors or mismatched result generation as compared to domain area perspectives, the process restarts to initial step so as to provide accurate results. Accuracy means the percentage of test set samples that are correctly classified by the classifier. Finally, visualization and knowledge representation are used to present the mined knowledge to the users and stored as new knowledge in the knowledge base. Incorporating the knowledge in to another system for implementation purpose, documentation and report for presenting the benefit of the knowledge to interested parties, incorporating the knowledge with previously known knowledge in the area are some of the important activities during this phase (Gheware et al., 2014).

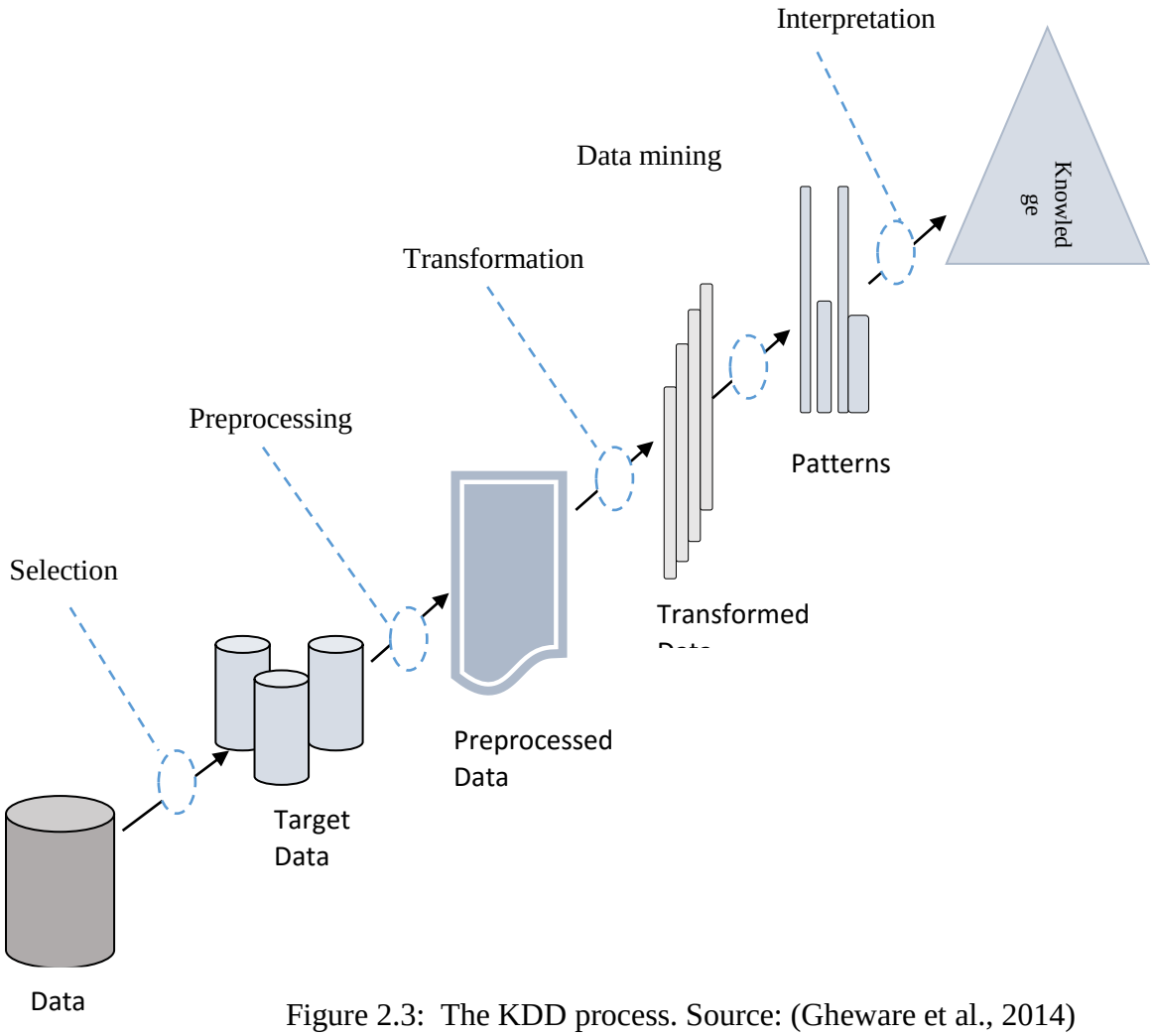


Figure 2.3: The KDD process. Source: (Gheware et al., 2014)

2.8.2. CRISP-DM PROCESS MODEL

The CRISP-DM (CRoss-Industry Standard Process for Data Mining) was first established inthe late 1990s by four companies: Integral Solutions Ltd. (a provider of commercial data miningsolutions), NCR (a database provider), DaimlerChrysler (an automobile manufacturer), and OHRA (an insurance company) (Cios et al., 2007). The CRISP-DM KDP model consists of six steps, which are summarized below.

Business understanding: This step focuses on the understanding of objectives and requirementsfrom a business perspective. It also converts these into a DM problem definition, and designs a preliminary project plan to achieve the objectives.

Data understanding: This step starts with initial data collection and familiarization with thedata. Specific aims include identification of data quality problems, initial insights into the data,and detection of interesting data subsets.

Data preparation: This step covers all activities needed to construct the final dataset, whichconstitutes the data that will be fed into DM tool(s) in the next step. It includes Table, record, attribute selection,data cleaning,construction of new attributes and transformation of data.

Modeling: At this point, various modeling techniques are selected and applied. Modeling usuallyinvolves the use of several methods for the same DM problem type and the calibration of theirparameters to optimal values. Since some methods may require a specific format for input data,often reiteration into the previous step is necessary.

Evaluation: After one or more models have been built that have high quality from a dataanalysis perspective, the model is evaluated from a business objective perspective. A reviewof the steps executed to construct the model is also performed. A key objective is to determinewhether any important business issues have not been sufficiently considered. At the end of this phase, a decision about the use of the DM results should be reached.

Deployment. Now the discovered knowledge must be organized and presented in a way that the customer can use. Depending on the requirements, this step can be as simple as generatinga report or as complex as implementing a repeatable KDP.

The model is characterized by an easy-to-understand vocabulary and good documentation. It divides all steps into substeps that provide all necessary details. It also acknowledgesthe strong

iterative nature of the process, with loops between several of the steps. In general,it is a very successful and extensively applied model, mainly due to its grounding in practical,industrial, real-world knowledge discovery experience (Cios et al., 2007).

The Major Applications of the CRISP-DM model has been used in domains such as medicine,engineering, marketing, and sales. It has also been incorporated into a commercial knowledgediscovery system called Clementine (Sundar, 2012).

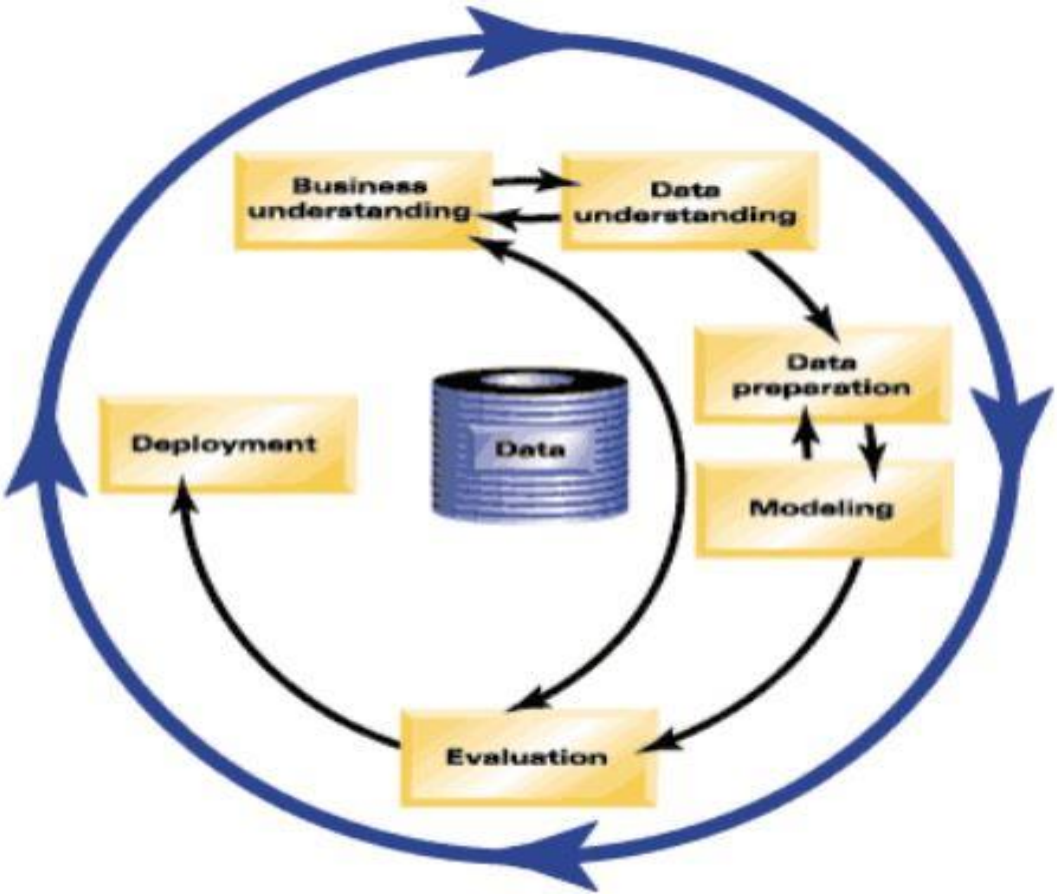


Figure 2.4: The CRISP-DM KD process model (Sundar, 2012).

2.8.3.HYBRID DM PROCESS MODELS

This model is known by models that combine aspects of both.It was developed based on the CRISP-DM model by adopting it toacademic research. The main differences and extensions include providing more general, research-oriented description of the steps and introducing a data mining step instead of the modeling step.Introducing several new explicit feedback

mechanisms, (the CRISP-DM model has only three major feedback sources, while the hybrid model has more detailed feedback mechanisms) and modification of the last step, since in the hybrid model, the knowledge discovered for a particular domain may be applied in other domains can listed as a diffrence between the two model (Cios et al., 2007).

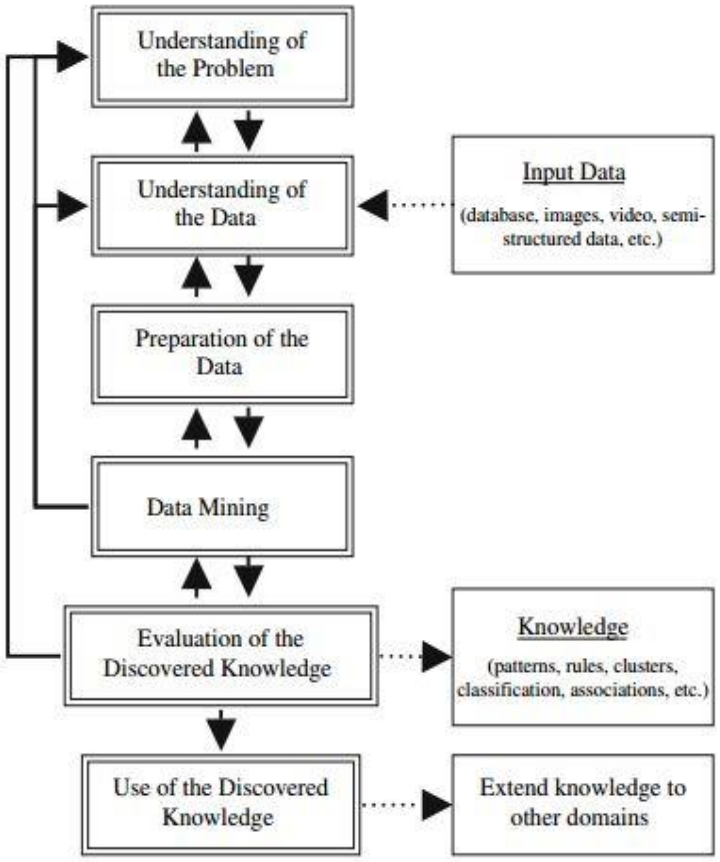


Figure 2.5: The six-step KDP model (Cios et al., 2007).

2.8.4. SEMMA

The SEMMA moethology is devoloped by the SAS Institute.The acronym SEMMA stands for sample, explore, modify, model, assess. Beginning with a statistically representative sample of your data, SEMMA intends to make it easy to apply exploratory statistical and visualization techniques, select and transform the most significant predictive variables, model the variables to predict outcomes, and finally confirm a model's accuracy (Obenshain et al., 2004).

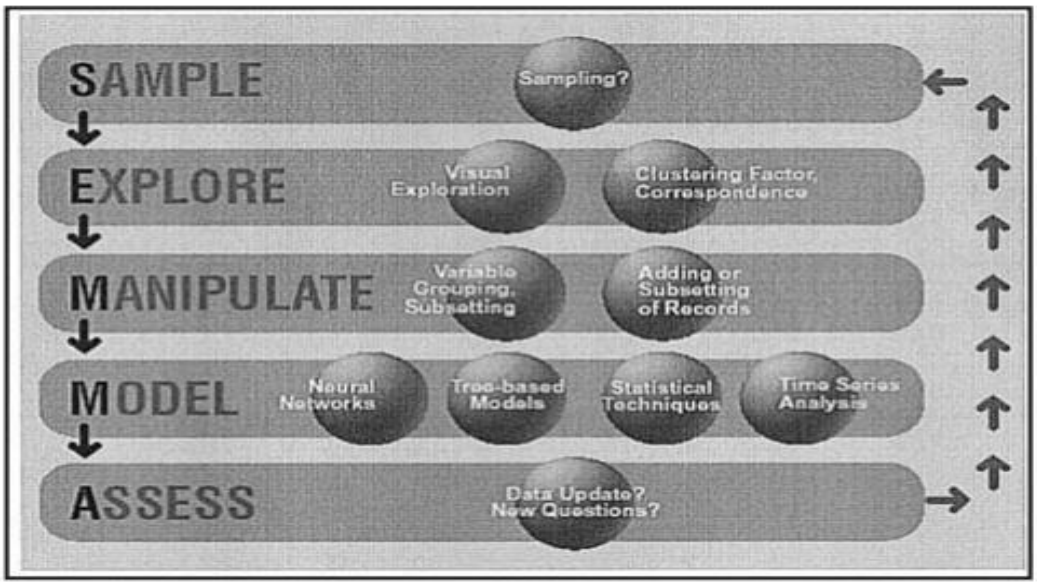


Figure 2.6: The SEMMA data mining process (obenshain et al., 2004).

The following steps are followed by SEMMA methodology (obenshain et al., 2004)

Step 1 (Sample):-This is where a portion of a large data set (big enough to contain the significant information yet small enough to manipulate quickly) is extracted.For optimal cost and computational performance, some (including the SAS Institute) advocates a sampling strategy, which applies a reliable, statistically representative sample of the full detail data. In the case of very large datasets, mining a representative sample instead of the whole volume may drastically reduce the processing time required to get crucial business information.

Step 2 (Explore):-This is where the user searched for unanticipated trends and anomalies in order to gain a better understanding of the data set. After sampling your data, the next step is to explore them visually or numerically for inherent trends or groupings. Exploration helps refine and redirect the discovery process.

Step 3 (Modify): This is where the user creates, selects, and transforms the variables upon which to focus the model construction process. Based on the discoveries in the exploration phase, one may need to manipulate data to include information such as the grouping of

customers and significant subgroups, or to introduce new variables. It may also be necessary to look for outliers and reduce the number of variables, to narrow them down to the most significant ones. One may also need to modify data when the "mined" data change. Because data mining is a dynamic, iterative process, you can update data mining methods or models when new information is available.

Step 4 (Model): This is where the user searches for a variable combination that reliably predicts a desired outcome. Once you prepare your data, you are ready to construct models that explain patterns in the data. Modeling techniques in data mining include artificial neural networks, decision trees, rough set analysis, support vector machines, logistic models, and other statistical models such as time series analysis, memory-based reasoning, and principal component analysis.

Step 5 (Assess): This is where the user evaluates the usefulness and the reliability of findings from the data mining process. In this final step of the data mining process user assesses the models to estimate how well it performs. A common means of assessing a model is to apply it to a portion of data set put aside (and not used during the model building) during the sampling stage. If the model is valid, it should work for thisreserved sample as well as for the sample used to constructthe model. Similarly, you can test the model against known data.

The SEMMA approach is completely compatible with the CRISP approach. Both aid the knowledge discovery process. Once models are obtained and tested, they can then be deployed to gain value with respect to business or research application (Delen, 2008).Summary of correspondences between KDD, SEMMA, CRISP-DM, AND HYBRID DM processes models are presented in Table 2.1 (Ana, 2008).

Table 2.1: Summary of data mining models

KDD	***SEMMA***	***CRISP-DM***	***HYBRID***
Pre KDD	----------	Business understanding	Problem domain understanding
Selection	Sample	Data understanding	Data understanding
Preprocessing	Explore		
Transformation	Modify	Data preparation	Data Preparation
Data mining	Model	Modeling	Data mining
Interpretation/evaluation of the discovered knowledge	Assessment	Evaluation	Evaluation
Post KDD	----------	Deployment of discovered knowledge	Use of discovered knowledge

From the Table 2.1 by doing a comparison of the models, some of them follow same steps to discovery process while others follow different steps. For example, in KDD and SEMMA stages the first approach is equivalent. Sample can be identified with Selection; Explore can be identified with Preprocessing; Modify can be identified with Transformation; Model can be identified with Data Mining; Assess can be identified with Interpretation/Evaluation. The researcher can conclude from the table shortly why hybrid is more convenient model that he will use to solve the stated problem, for this Crisp and Hybrid model is ideally related but at first the business understanding and problem domain understanding is their difference for this reason the researcher will still strongly agree with hybrid since the problem domain goes well with the maternal mortality. On here the hospital never needs the business area of maternal mortality rate even though it is important to minimize the problem part. Beside this, the maternity ward will like to use the knowledge for the problem after use and distribution of that knowledge is preferable after they applied the discovered knowledge purposely.

2.9. THE FUNCTIONS OF DATA MINING

There are two main functions of data mining. Those are classification and clustering. By classifying and clustering we can manage the raw data from data ware house, data base and other institutional information resources.

2.9.1. CLASSIFICATION

Classification is used to classify data into predefined categorical class labels. Class in classification, is the attribute or feature in a data set, in which users are most interested. It is defined as the dependent variable in statistics. To classify data (or records), a classification algorithm creates a classification model consisting of classification rules. Classification is a task that occurs very frequently in everyday life. Essentially it involves dividing up objects so that each is assigned to one of a number of mutually exhaustive and exclusive categories known as classes. The term 'mutually exhaustive and exclusive' simply means that each object must be assigned to precisely one class, i.e. never to more than one and never to no class at all (Srideivanai ,2014).

2.9.2. CLUSTERING

Clustering techniques apply when there is no class to be predicted but the instances are to be divided into natural groups. These clusters presumably reflect some mechanism that is at work in the domain from which instances are drawn, a mechanism that causes some instances to bear a stronger resemblance to each other than they due to the remaining instances. Clustering naturally requires different techniques to the classification and association learning method. There are different ways in which the result of clustering can be expressed. The groups that are identified may be exclusive: Any instance belongs in only one group or they may be overlapping: An instance may fall into several groups. An instance belongs to each group with a certain probability or they may be hierarchical: A rough division of instances into groups at the top level and each group refined further perhaps all the way down to individual instances. Really, the choice among these possibilities should be dictated by the nature of the mechanisms that are thought to underlie the particular clustering phenomenon (Kaur et al., 2013).

2.10. WHAT IS ASSOCIATION RULES?

Given a collection of frequent item sets F, to generate association rules we iterate over all item sets Z ∈ F, and calculate the confidence of various rules that can be derived from the item set. Formally, given a frequent item set Z ∈ F, we look at all proper subsets X ⊂ Z to compute rules of the form (Delen, 2008).

$$X \xrightarrow{s,c} Y, \text{ where } Y = Z \setminus X$$

where $Z \setminus X = Z - X$. The rule must be frequent since

$$s = sup(XY) = sup(Z) \geq minsup$$

Thus, we have to only check whether the rule confidence satisfies the minconf threshold. We compute the confidence as follows.

$$c = \frac{sup(X \cup Y)}{sup(X)} = \frac{sup(Z)}{sup(X)}$$

```
AssociationRules (F, minconf):
foreach Z ∈ F, such that |Z| ≥ 2 do
    A ← {X | X ⊂ Z, X ≠ ∅}
    while A ≠ ∅ do
        X ← maximal element in A
        A ← A \ X // remove X from A
        c ← sup(Z)/sup(X)
        if c ≥ minconf then
            print X ⟶ Y, sup(Z), c
        else
            A ← A \ {W | W ⊂ X} // remove all subsets of X from A
```

2.11. DATA MINING ALGORITHM

The rule (algorithm) is used to create models with predictive capabilities. It provides new ways of exploring and understanding data. It learns from the "evidence" by calculating the correlation between the target (i.e., dependent) and other (i.e., independent) variables in Data Mining one of the most common tasks is to build models for the prediction of the class of an object on the basis of its attribute. Here the object can be seen as a customer, patient, transaction, e-mail message or even a single character. Attributes of such objects can be, for example for the patient object, hearth rate, blood pressure, weight and gender, whereas the class of the patient object would most commonly be positive/negative for a certain disease. Before applying data mining algorithms to maternal mortality data, researchers must understand what kind of data mining algorithms exists and how they function (Aditya, 2012).

Generally, data mining algorithms are classified into two categories: descriptive (or unsupervised learning) and predictive (or supervised learning). Descriptive data mining clusters data by measuring the similarity between objects (or records) and discovers unknown patterns or relationships in data so that users can readily understand a huge amount of data. Descriptive data mining is exploratory in nature. This type of data mining includes clustering, association, summarization, and sequence discovery. Prediction data mining infers prediction rules (a.k.a. classification/prediction models) from (training) data and applies the rules to unpredicted/unclassified data. Prediction data mining includes classification, regression, time series analysis, and prediction (Ian et al ., 2011).

2.11.1. DECISION TREES

Decision trees are a class of data mining techniques that have roots in traditional statistical disciplines such as linear regression. Decision trees also share roots in the same field of cognitive science that produced neural networks. Decision trees are a simple, but powerful form of multiple variable analyses. A decision tree is a classifier expressed as a recursive partition of the instance space. The decision tree consists of nodes that form a rooted tree, meaning it is a directed tree with a node called "root" that has no incoming edges. All other nodes have exactly one incoming edge. A node with outgoing edges is called an internal or test node. All other nodes are called leaves (also known as terminal or decision nodes). In a decision tree, each internal node splits the instance space into two or more subspaces according to a certain discrete

function of the input attributes values. In the simplest and most frequent case, each test considers a single attribute, such that the instance space is partitioned according to the attribute's value. In the case of numeric attributes, the condition refers to a range (Petri, 2010).

2.11.2. J48 DECISION TREE

J48 classifier is among the most popular and powerful decision tree classifiers. C5.0 and J48 are the improved versions of C4.5 algorithms. J48 is an open source implementation of C4.5 algorithms in weka. J48 adopts a greedy approach in which decision tree is constructed in a top down recursive divide and conquer manner. The decision tree algorithm works as follows:

Create a node N;
if tuples in D are all of the same class, C then

return N as a leaf node labeled with the class C;

if attribute list is empty

then return N as a leaf node labeled with the majority class in D; //
majority voting
apply Attribute selection method(D, attribute list) to find the "best"
splitting criterion;
label node N with splitting criterion;
if splitting attribute is discrete-valued and multiway splits allowed
then // not restricted to binary trees
attribute list attribute list _ splitting attribute; // remove splitting
attribute
for each outcome j of splitting criterion // partition the tuples and
grow subtrees for each partition
let Dj be the set of data tuples in D satisfying outcome j; // a partition
if Dj is empty then (
attach a leaf labeled with the majority class in D to node N;
else attach the node returned by Generate decision tree(Dj, attribute

list) to node N; endfor

return N;. (Kaur.et al, 2013)

Kumar and Sahoo (2012), also stated that J48 are the improved versions of C4.5 algorithms or can be called as optimized implementation of the C4.5. The output of J48 is the Decision tree.

2.11.3. NAIVE BAYES

Naïve Bayes is statistical algorithm that is based on Bayesian theorem. A Bayesian classifier tries to estimate the conditional probability that an input document belongs to a category. It compares "data in a document d" to "text that would be generated by the model associated with a category c." Then it computes an estimate of the likelihood that d belongs to c.

In text categorization, NB calculates probability values in order to assign category labels. Firstly, prior category probabilities are calculated. P (c_i) is prior probability that document d_i is in c_i if we knew nothing about "the text in d_i." Then we multiply it with the probability that d_i is generated by c_i. The result is called the posterior probability P (c_i| d_i), which can be computed from the product of the prior probability P (c_i) and the likelihood P (d_i | c_i) according to Bayes theorem:

$$P(c_i \mid d_i) = \frac{P(di|ci)\ P(ci)}{P(di)}$$

Since the probability that a document d_i occurs in the corpus, P (d_i), is a fixed value for a given document d_i, we do not need to estimate it. The estimation of the posterior probability P (c_i | d_i) is thus converted to the estimation of the prior probability P (c_i) and the likelihood P (d_i | c_i). If the terms of the input document are assumed to be conditionally independent given the category, the likelihood P (d_i | c_i) can be simply calculated by multiplying the likelihood of category c_i with respect to each term:

$$P(d_i \mid c_i) = \prod_{k=1}^{|T|} P(\mathrm{tk} \mid \mathrm{ci})$$

Where t_k is the weight of the k^{th} term in document d_i, and $|T|$ is the total number of terms. The probability distributions $p(c_i)$ and $P(t_k \mid c_i)$ are usually assumed to have known parametric forms, and the learning task is essentially the estimation of the parameters.

2.11.4. PART RULE INDUCTION

PART (Partial Decision Tree) is a rule induction algorithm which grabs rule from a decision tree. A partial decision tree is an ordinary decision tree that contains branches to undefined sub trees (Ian, 2005). To generate such a tree, the construction and pruning operations are integrated in order to find a "stable" sub tree that can be simplified no further (Ian, 2005). Once this sub tree has been found, tree building ceases and a single rule is read off. The following algorithm depicts the steps and procedures followed in implementing PART rule induction.

Initialize E to the instance set

For each class C, from smallest to largest

BUILD:

Split E into Growing and Pruning sets in the ratio 2:1

Repeat until (a) there are no more uncovered examples of C; or (b) the description length (DL) of rule set and examples is 64 bits greater than the smallest DL found so far, or (c) the error rate exceeds 50%:

GROW phase:

Grow a rule by greedily adding conditions until the rule is 100%accurate by testing every possible value of each attribute and selecting the condition with greatest information gain G

PRUNE phase: Prune conditions in last-to-first order. Continue as long as the worth W of the rule increases

OPTIMIZE:

GENERATE VARIANTS:

For each rule R for class C,

Split E afresh into Growing and Pruning sets

Remove all instances from the Pruning set that are covered by other rules for C

Use GROWS and PRUNE to generate and prune two competing rules from the newly split data:

R1 is a new rule, rebuilt from scratch;

R2 is generated by greedily adding antecedents to R.

Prune using the metric A (instead of W) on this reduced data

SELECT REPRESENTATIVE:

Replace R by whichever of R, R1 and R2 has the smallest DL.

MOP UP:

If there are residual uncovered instances of class C, return to the BUILD stage to generate more rules based on these

CLEAN UP:

Calculate DL for the whole rule set and for the rule set with each rule in turn omitted; delete any rule that increases the DL Remove instances covered by the rules just generated.

Continue….

2.12. DATA MINING APPLICATION IN HEALTH CARE

Data mining techniques has been used intensively and extensively by many organizations. In healthcare, data mining is gradually increasing popularity, if not by any case, becoming increasingly essential. Data mining applications can greatly benefit all parties involved in the healthcare industry.Several factors have motivated the use of data mining applications in healthcare.it is defiend that data mining is the way of mining precious knowledge from a bunch of data or recoreds,in today era there many records in the hospitals or health sectors.due to this many health organizations needs the application of data mining techniques in their records to find the hidden patterns from their data (Pradhan, 2014).

Data mining solved different types of health sector problem by using different types of model development.this means shorltly predicting the future contents from the known value and

unknown value. Data mining can be defined as the process of finding previously unknown patterns and trends in databases and using that information to build predictive models.As the all data mining applied in medical decision support,diagnosis and treatment,healthcare resources management,Prediction of inpatient length of stay and Unhealthy insurance practices (David, 2013).

2.13. RELATED WORKS

Data mining solves many problems by using different types techniques. It is unforgettable that data mining is known by finding a precious knowledge. Data mining is a tool which is works as interdisciplinary for other related field. These are because of so many researchers use this tool as a basic tool and solve the real world problem.

Aneeshkumar & Venkateswaran (2012) explains the art of data mining tool in predicting the Chances of Liver Disease in Ectopic Pregnant Groups, they used the regression method and K fold cross validation to subgroup the data of the for liver disease. On here they used the data mining with the statistical method. Medical data mining techniques like Association Rule Mining, Clustering, and Classification Algorithms such as Decision tree, C4.5 Algorithm are implemented to analyze the different kinds of heart based problems. C4.5 Algorithm and Clustering Algorithm like K-Means are the data mining techniques used in medical field. With the help of this technique, the accuracy of disease can be validated. Classification is one of the supervised learning methods to extract models describing important classes of data. Three classifiers Decision Tree, Naïve Bayes and Classification via clustering have been used to diagnose the Presence of heart disease in patients (Nishara & Gomathy, 2013).

Zenebe (2013), PART pruned rule induction model, J48 un pruned tree and Navies Bayes are appeared with good predictive performance for nutritional status of under-five children. From all the scenarios experimented, all models reveal the better performance in predicting True positive cases or sensitivity; than predictive performance of True negative case or specificity.As sensitivity and specificity has greater importance than general accuracy of the classifier in clinical and medical fields, models are better compared based on WROC area. The model created using PART pruned rule induction classifier registers good performance (i.e. 97.8% WROC area) and hence selected for further analysis/rule tracing. PART is a good algorithm to predict the under five-year nutrient ion

Dursun et al. (2004) also used (artificial neural networks and decision trees) to predict breast cancer survivability. Additionally, they used 10-fold cross-validation methods to measure the unbiased estimate of the three prediction models for performance comparison purposes. In order to perform the research reported in this manuscript, we used the data contained in the SEER Cancer Incidence Public-Use Database for the years 1973-2000. After the data preprocesses data cleansing and data preparation strategies, the final dataset, which consisted of 17 variables (16 predictor variables and 1 dependent variable) and 202,932 records, was constructed. The c=accuracy measurement was measured by stratified 10-fold cross validation and the result also obtained by the result testing on the data set. the result shows the ANN model achieved a classification accuracy of 0.9121 with a sensitivity of 0.9437 and a specificity of 0.8748. The logistic regression model achieved a classification accuracy of 0.8920 with a sensitivity of 0.9017 and a specificity of 0.8786. However, the decision tree (C5) preformed the best of the three models evaluated. The decision tree (C5) achieved a classification accuracy of 0.9362 with a sensitivity of 0.9602 and a specificity of 0.9066. For each fold of each model type, the detailed prediction results of the validation datasets are presented in form of confusion matrixes. A confusion matrix is a matrix representation of the classification results.

Shegaw (2002) also used best performing neural network model and decision tree classifier for the prediction of child mortality pattern. The methodology employed consisted of three basic steps; data collection, data preparation, and model building and testing. However, since a data mining task is an iterative process, these steps were not followed strictly in linear order. In order to build models that can predict the risk of child mortality, several models were built by employing both neural network and decision tree approaches. The best performing neural network model and decision tree classifier were then chosen and evaluated using ten previously unseen records of children. Using the neural network approach, the best model was identified for the training made by using the default parameters (i. e. training tolerance of 0.1, learning rate of 1.0, and smoothing factor of 0.9) and the following 9 input variables: “ENVIRN”, “AGE”, “OUTMIG”, “HHRELIG”, “HHETHNIC”, “HHLITERAC”, “HHHEALHTH”, “HHWATER”, AND “WINDOWS”. This model had an accuracy rate of 93% (classified 102 of the 110 test cases correct) at a testing tolerance of 0.4 and was tested with accuracy of 88 % (classified 97 of the 110 test cases correct) at testing tolerances of 0.2 and 0.1.

Shegaw (2002) , also stated that encouraging results were obtained by employing both neural networks and decision tree approaches. Although both neural network and decision trees showed comparable accuracy and performance in predicting the risk of child mortality, the decision tree approach seems more applicable and appropriate to the problem domain since it provides additional features such as simple and easily understandable rules that can be used by non-technical health care professionals as well as health care planners and policy maker.

Lijia Guo and Morgan Wang (2003), used the popular data mining techniques include Bayesian analysis, neural networks, genetic algorithms, decision trees and logistic regression. The propose to use a hybrid method that combines the strength of both logistic regression and decision trees in the paper. The proposed hybrid method includes three steps. The first step is to identify the importance of the risk factors in determining the advanced age mortality distribution. We fit a logistic regression model using only age and square term of age as the depend variables because it is well knowing that mortality rate is grow exponentially with age. If the mortality rate only depends on age and square term of age, the residuals of this logistic regression model should not have any special pattern. If there are other variables that need to include in the model, the decision trees should be able to identify them through these residuals. Thus, the next step of this hybrid method is using decision trees on these residuals. In this study, the decision tree algorithm identifies six segments and picks up different set of variables and interaction terms in each segment. Finally, a logistic regression model for each segment is developed. All the important variables and important interaction terms are included in each segment logistic regression model. Since the number of records available in each segment is different, important variables and interaction terms selected in each segment are different in this study. Consequently, the predicted power for each segment logistic regression model is different. Detail description of each segment's logistic regression model is given in the section of the paper. Generally, from these literatures review the researcher will gain one useful information to solve the stated problem, these literature review will cover the rough starting for the stated problem. From the reviewed literature different types researcher will prove the techniques and algorithm that the researcher will use in the future. Without any doubt the listed algorithm and techniques that the researcher raises under chapter one of this proposal PART tree, J48 and naïve Bayesian will solve the problem stated under chapter one.

Temesgen (2015), experimented that classification with both decision tree and neural network for the application of data mining techniques to discover cause of under five children admission to pediatric ward and stated that data mining technique is applicable on pediatric dataset in developing a model that support the discovery of the causes of under-five children admission to pediatric ward. He got the encouraging result with J48 algorithm. The decision tree algorithm J48 has higher accuracy (94.77%), weighted true positive rate (94.7%), weighted false positive rate (5.3%), weighted receiver operating characteristics curve (0.99) and performs much faster than multilayer perceptron. In addition to that, models built by using neural network, were incomprehensible for a human and the extraction of business knowledge from it was found to be difficult. According to interesting rules in J48, presenting complaint of not taking any food, fluid or breast feeding (98.32%), low weight for age without sunken eyes (92.31%) and very low weight for age but not in association with restless or irritable (98.33%) were among the cause of under-five children admission to pediatric ward without any consideration of health information management system admission disease classification criteria.

Tefera et al. (2014) the six-step hybrid knowledge discovery process model is used as a framework for 15961 instances that have undergone urinary fistula repair in Addis Ababa Fistula Hospital are used for both predictive association rule extraction and predictive model building. Apriori algorithm is used to extract association rules while classification algorithms J48, PART, Naïve Bayes and multinomial logistic regression are used to build predictive models. After they applied those three algorithms in the study they concluded Predictive association rules from Apriori have shown frequent co-occurrence of less severity of injury with cured outcome. The predictive model from PART-M2-C0.05-Q1 scheme has shown an area under WROC curve of 0.742. Area under the ROC curve for residual outcome (ROC Residual=0.822) from this algorithm is better than Naïve Bayes and logistic, while the areas under the ROC curves for the other outcomes are greater than the model from J48 Predictive association rules from Apriori have shown frequent co-occurrence of less severity of injury with cured outcome. The predictive model from PART-M2-C0.05-Q1 scheme has shown an area under WROC curve of 0.742. Area under the ROC curve for residual outcome (ROC Residual=0.822) from this algorithm is better than Naïve Bayes and logistic, while the areas under the ROC curves for the other outcomes are greater than the model from J48.Lastly Predictive model is developed with the use of PART-M2-C0.05-Q1. The predictive association

rules and predictive model built with the use of data mining techniques can assist in predicting urinary fistula surgical repair outcome.

Getachew (2013) also used association rules mining using Apriori algorithm and three commonly used and popular classification algorithms (J48, Naïve Bayes and PART) and long process of data cleansing, data and dimensionality reduction and transformation used it to build the association and prediction models on 11,440 instances and 10 attribute of ART data set from Addama and Ambo hospital for the prediction of ART. Also other data mining algorithm was stated in his work but he proved the J48, Naïve Bayes and PART are the good algorithm for the prediction of ART.

Selam (2011) also used the hybrid model approach for the prediction of occurrence of measles outbreak in Ethiopia. He got the result by using, J48 decision tree and Naïve Bayes but the first algorithm more predictor than Naive Bayes. Hence J48 decision tree with 97.06% accuracy prediction model building was selected to extract interesting rules to cite.

Tesfahun (2012) also tried to predict the adult mortality by using data mining techniques and he got the encouraging result to develop a model. During his study he created a model which predictive accuracy result of 97.2% and 98.5% correctly predictive performance of individual as alive cases indeed they are alive. The best performing decision tree model was then chosen and evaluated using previously unseen records of adult.

Teketel (2013) used the hybrid approach model to predict and create a model by using data mining techniques to minimize the occurrence of tuberculosis. For this model he used MLP, J48 and SMO for the model. On this study the best performing algorithms are J48 classifier followed by MLP classifier and SMO classifier. The least performed algorithm was Naïve Bayes classifier. The best selected model in this study was generated by J48 decision tree with all attributes and accuracy of this model is 95.24%.

According to the above cited researchers, the evaluation of best performed algorithms compared based on accuracy, sensitivity, specificity, false prediction rate and time taken to build model. They also used the 10-fold cross validation to evaluate their performance of the model. Additionally, they used the J48 decision tree, Naïve Bayes and PART and they got encouraging result on the prediction and model creation. According to the researcher knowledge J48 decision

tree and Naïve Bayes algorithm used at one place sometimes and adding one algorithm beside J48 decision tree and Naïve Bayes will produce a good result for the purpose of prediction techniques in health area. Adding one algorithm and above on the J48 decision tree and Naïve Bayes is the gap and recommendation that many researchers left in their paper and WEKA tool is used as primary.

But they used the old version of WEKA, for this reason the advanced version of WEKA and other un used algorithm will be one of the gap that the researcher understands. Also applying the data mining techniques will provide a good model and result. Therefore, decision tree, Naïve Bayesian and PART are used for the purpose of classification and prediction of events in the future with useful knowledge. Using those algorithms and adding hybrid model approach would success the work of maternal mortality rate prediction. Additionally, using the advanced version of WEKA with modified algorithm will produce a good result for the maternal mortality in maternity ward.

To sum up the related work section, different researchers used PART rule induction, J48 decision tree and Naïve Bayes including neural network for the purpose of prediction. But selecting the gap that is un touched by the researchers which is predicting maternal mortality for pregnant women which determine their future status would make this study different. Good performance and accuracy is deserved according to the researcher's objectives but not for maternal mortality. This only covers the pregnant woman which comes to maternity ward for the purpose of labor and delivery. Finding the attribute which is approved by the three selected model would make this study different from the above related work.

CHAPTER THREE

DATA MINING ARCHITECTURE AND METHODS

In this study, an attempt has been made to apply data mining techniques in the health care sector in particular for predicting maternal mortality rate.

3.2. THE ARCHITECTURE OF THE SYSTEM

It is important to imagined or abstract about a specific problem before the researcher starts to dig out. The researcher is being mentioned about the methodology and tools which is used in maternal mortality for prediction purpose by the help of data mining techniques. The architecture of the problem with respect to the methodology of data mining can be designed for purpose of predicting maternal mortality. Chaudhary (2015) explains the Data mining architecture contains user interface,pattern eveluation,data mining engine,data base,data cleaning consilidation and selection. User interface module interacts between user and data exploring system. It allows the subscriber to do interaction with the system by explaining his query and simultaneously by identifying information in order to help in search and to carry out exploratory data mining based on the intermediate data mining results.Therefore, the researcher takes the architecture of data mining that is raised above and develops the down figured logical view of the main stated problem under chapter one of this study.

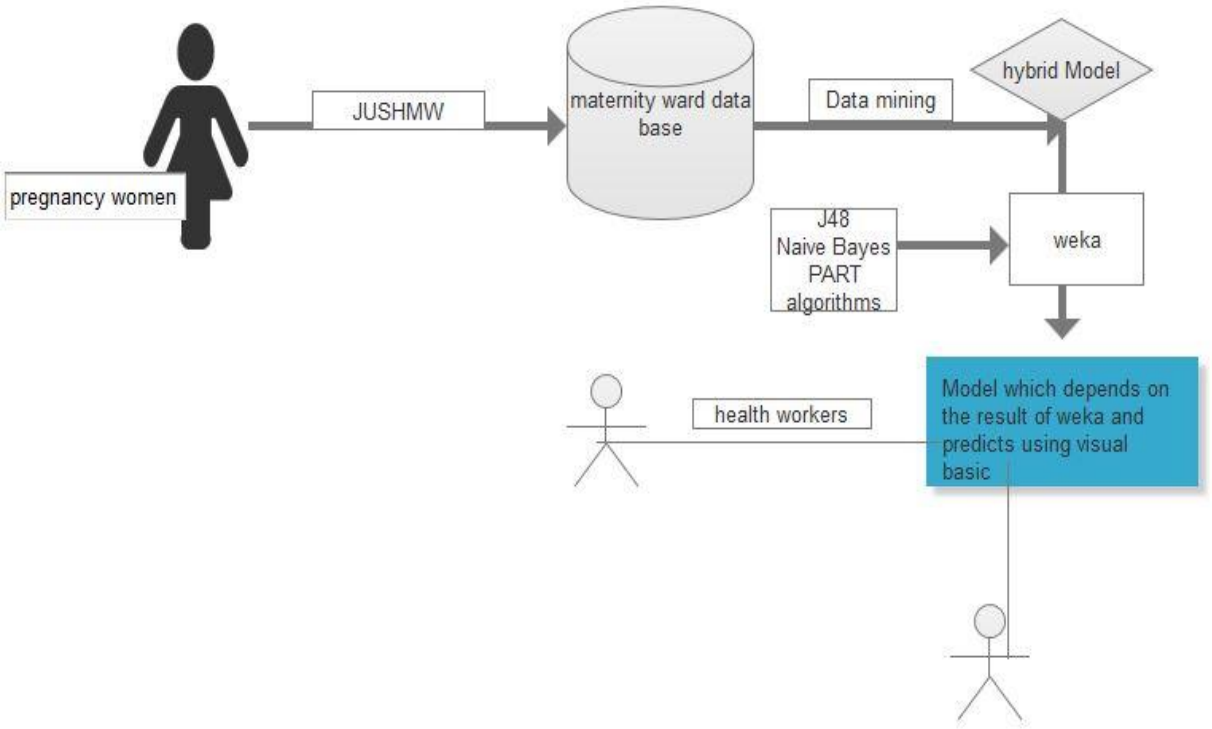


Figure 3.1: The logical view of the system.

3.3. DESCRIPTIVE AND PREDICTIVE TASKS OF DATA MINING.

Descriptive and predictive tasks are the two main data mining functionality in finding a precious knowledge. The Descriptive tasks present the general properties of data stored in database. This means with the concept of this study from the maternity ward data which attributes has a relationship with other and how they performed with the internal class attributes of pregnant women. The general attributes which predict the future life status of pregnant women's life but not still un identified weather they could predict or not. Since the data stored contains the general attributes for the maternal mortality. The descriptive tasks are used to find out patterns in data i.e. cluster, correlation, trends and anomalies. Predictive data mining tasks predict the value of one attribute on the bases of values of other attributes, which is known as target or dependent variable and the attributes used for making the prediction are known as independent variables (Chaudhary, 2015). Descriptive describes the data set in a concise and summary manner and presents interesting general properties of the data. Predictive constructs one or a set of models, performs infrence on the available set of data and attempts to predict the behavior of a new data set. As a general the predive model constructs a set of models which eaither

supported by J48 decision tree or Naïve Bayes or PART rule induction by selcting the independent variables from the maternity ward data. Then by selecting the most performing model from the constructed internaly independent variables and classifying according to the internal relationship of attributes of pregnant women is possible. Depend on those out put of the model future status of pregnant woman would predict.

3.4. TECHNIQUES OF DATA MINING

3.4.1 CLASSIFICATION

Classification is one of the most popularly used methods of Data Mining in Healthcare sector. It divides data samples into target classes. It is a supervised learning approach having known class categories. Binning and multilevel are the two methods of classification. In binary classification, only two possible classes such as, “high” or “low” risk patient may be considered while the multiclass approach has more than two targets for example, “high”, “medium” and “low” risk patient. Data set is partitioned as training and testing dataset. It consists of predicting a certain outcome based on a given input. Training set is the algorithm which consists of a set of attributes in order to predict the outcome. In order to predict the outcome it attempts to discover the relationship between attributes. Goal or prediction is its outcome. There is another algorithm known as prediction set. It consists of same set of attributes as that of training set. But in prediction set, prediction attribute is yet to be known. In order to process the prediction it mainly analyses the input. The term which defines how “good” the algorithm is its accuracy (Parvez, 2015).

In the supervised classification, there is mapping of input data set to finite set of discrete class labels. Input data set X € Ri , where i is the input space dimensionally and discrete class label Y € 1.....T, where T is the total number of class types. And this is modeled in the term of equation Y=Y(x, w), w is the vector of adjustable parameters. Classification techniques in data mining can be as follows (Goel, 2015).

Decision tree induction: From the class labeled tuples the decision tree is build. Decision tree is tree like structure in which there are internal node, branch and leaf node. Internal node specifies the test on attribute, branch represents the outcome of the test and leaf node represents the class label. Two steps that are learning and testing are simple and fast. The main goal is to predict the output for continuous attribute but decision tree is less appropriate

for estimating tasks. There may be errors in predicting the classes by using decision tree approach. Pruning algorithms are expensive and building decision tree is also an expensive task as at each level there is splitting of node.

Rule – Based classification: It is represented by set of IF- THEN rules. First of all how many of these rules are examined and next care is about how these rules are build and can be generated from decision tree or it may be generated from training data using sequential covering algorithm. Expression for rule is:

IF condition THEN conclusion

Now we define accuracy and coverage of S by following expression

Coverage (R) = $Ntotal \ / \ I\boldsymbol{D}I$

Coverage (R) = $Ncorrect \ / \ Ntotal$

Classification prediction encompasses two levels: classifier construction and the usage of the classifier constructed. The former is concerned with the building of a classification model by describing a set of predetermined classes from a training set as a result of learning from that dataset. Each sample in the training set is assumed to belong to a predefined class, as determined by the class attribute label. The model is represented as classification rules, decision trees, or mathematical formula. The later involves the use of a classifier built to predict or classify unknown objects based on the patterns observed in the training set (Danso, 2006).

3.4.2. MEASURING CLASSIFIER PERFORMANCE

The evaluation measures in classification problems are defined from a matrix with the numbers of examples correctly and incorrectly classified for each class, named confusion matrix.The confusion matrix for a binary classification problem (which has only two classes – positive and negative). The confusion matrix is used as an indication of the properties of a classification (discriminant) rule. It contains the number of elements that have been correctly or incorrectly classified for each class. On the diagonal the number of observations that have been correctly classified for each class; the off-diagonal elements indicate the number of observations that have been incorrectly classified or helps to see if the system is confusing two classes (Maimon, 2008).

Table 3.1: Confusion Matrix

	Predicted class	
True class	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Based on the above Table 3.1, FP, FN, TP and TN can be described, False positives (FP) is examples predicted as positive, which are from the negative class. A false negative (FN) is examples predicted as negative, whose true class is positive. A true positive (TP) is examples correctly predicted as pertaining to the positive class. A true negative (TN) is examples correctly predicted as belonging to the negative class. We can calculate the accuracy from the above table by the following formula (Witten, 2005).

$$Acc = \frac{|\mathrm{TN}|+|\mathrm{TP}|}{|\mathrm{FN}|+|\mathrm{FP}|+|\mathrm{TN}|+|\mathrm{TP}|} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1)$$

Accuracy and Error are general measures and can be directly adapted to multiclass classification problems. Shortly we can correlate both of them together.

$$Err = \frac{|\mathrm{FN}|+|\mathrm{FP}|}{|\mathrm{FN}|+|\mathrm{FP}|+|\mathrm{TN}|+|\mathrm{TP}|} = 1 - Acc \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.2)$$

The recall (R) and specificity (Spe) measures evaluate the effectiveness of a classifier for each class in the binary problem. The recall, also known as sensitivity or true positive rate, is the proportion of examples belonging to the positive class which were correctly predicted as positive. The specificity is the percentage of negative examples correctly predicted as negative.

$$R = \frac{|\mathrm{TP}|}{|\mathrm{TP}|+|\mathrm{FN}|} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.3)$$

$$Spe = \frac{|\mathrm{TN}|}{|\mathrm{FP}|+|\mathrm{TN}|} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.4)$$

Precision (P) is a measure which estimates the probability that a positive prediction is correct. We can calculate by the equation below and we can combine with recall by originating the F-measure.

$$P = \frac{|\mathrm{TP}|}{|\mathrm{TP}|+|\mathrm{FP}|} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.5)$$

$$F - Measure = \frac{(\beta^2+1)*\mathrm{P}*\mathrm{R}}{\beta^2*P+R} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.6)$$

3.4.3. ROC CURVE

Receiver Operating Characteristics (ROC) curve is a plot of TPR against FPR which depicts relative trade-offs between benefits (true positives) and costs (false positives). Receiver Operating Characteristics (ROC) graphs have long been used in signal detection theory to depict the tradeoff between hit rates and false alarm rates over noisy channel. ROC curves also offer a more complete picture of the performance of the classifier. To compare classifiers, we may want to reduce ROC performance to a single scalar value representing expected performance and common method is to calculate the area under the ROC curve, abbreviated AUC. AU C(h)>AU C(g): classifier h has better average performance (Provost and Fawcett , 1998).

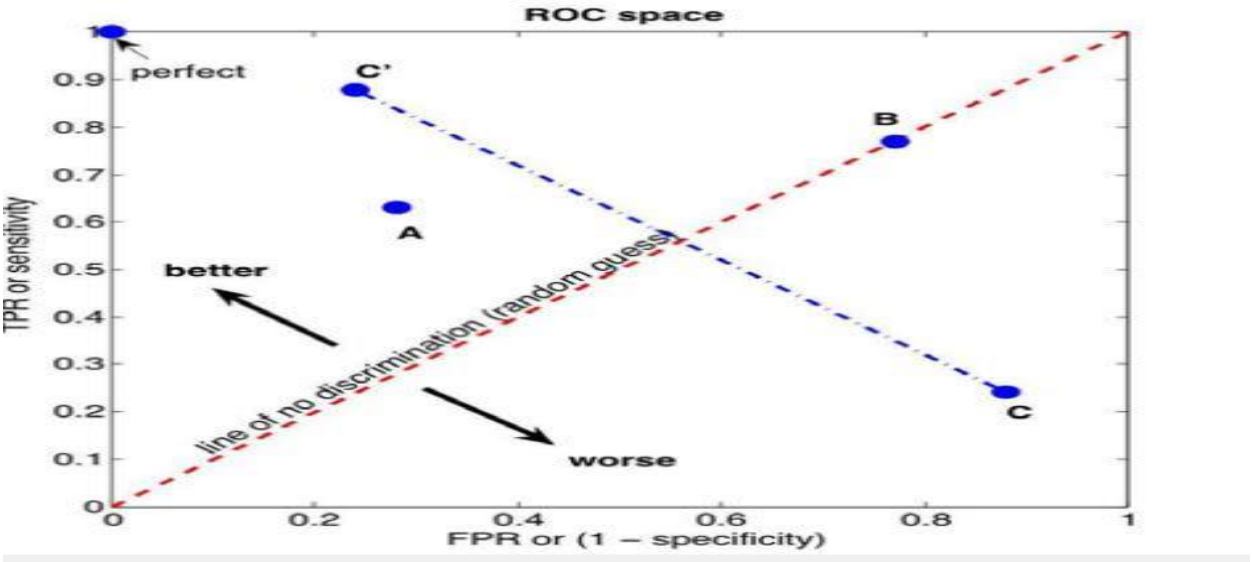


Figure 3.2: ROC curve graph (Provost and Fawcett, 1998).

3.5. PREDICTION

Numeric prediction is the task of predicting continuous or ordered values for given input. This prediction is most known by numeric prediction wich is almost known by regression. Regression is a statistical methodology that was developed by Sir Frances Galton (1822–1911). Regression and prediction is synonymously used by different writers, therefore regression widely used with prediction. Regression analysis can be used to model the relationship between one or more independent or predictor variables and a dependent or response variable (which is continuous-valued). In the context of data mining, the predictor variables are the attributes of interest describing the tuple (i.e., making up the attribute vector). In general, the values of the predictor variables are known (Jiawei & Micheline, 2006).

3.6. LOGISTIC REGRESSION

Logistic Regression is a type of regression model where the dependent variable (target) has just two values, such as: 0, 1 or Y, N and F, T. Logistic regression is an approach to prediction. Sometimes used for the prediction of dichotomous outcome. When using the logistic distribution, we need to make an algebraic conversion to arrive at our usual linear regression equation (which we have written as Y = B0 + B1X + e) (Sayad, 2010).

$$P = \frac{1}{1+e^{-(\beta_{0+}\beta iX)}} = \ln\left(\frac{p}{1-p}\right) = \beta 0 + \beta 1X \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.7)$$

The logistic distribution constrains the estimated probabilities to lie between 0 and 1. Maximum Likelihood Estimation is a statistical method for estimating the coefficients of a model. The regression coefficients measure the predictive capability of the independent variables. Exp (β) is the odds ratio corresponding to a one unit change in X.

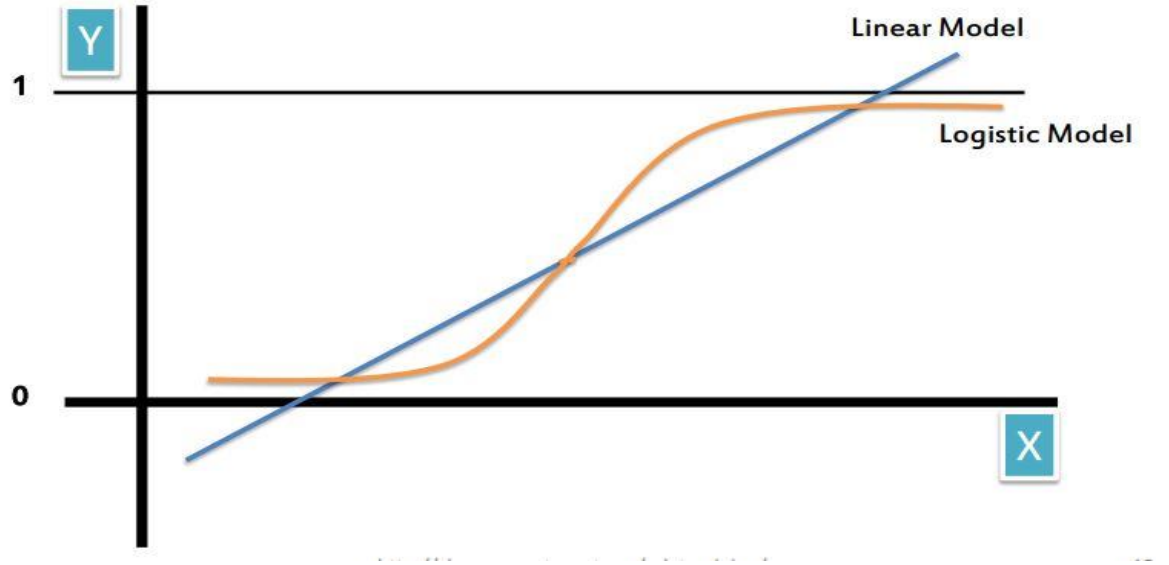


Figure 3.3: logistic regression model (Saed Sayad, 2010): retrieved from: http://chem-eng.utoronto.ca/datamining).

3.7. OUTLIER ANALYSIS

Outlier is data objects that do not comply with the general behavior or model of the data or data objects, which are grossly different from or inconsistent with the remaining set of data. Outliers can be caused by measurement or execution error. Many data mining algorithms try to minimize the influence of outliers or eliminate them all together. Outlier mining can be described as follows: Given a set of n data points or objects and k, the expected number of outliers, find the top k objects that are considerably dissimilar, exceptional, or inconsistent with respect to the remaining data. The outlier mining problem can be viewed as two sub problems: define what data can be considered as inconsistent in a given data set and find an efficient method to mine the outliers so defined. The problem of defining outliers is nontrivial (Han, 2006).

3.8. ALGORITHMS AND METHODS

3.8.1. NAIVE BAYES ALGORITHM

The Naive Bayes algorithm is a classification algorithm based on Bayes rule and a set of conditional independence assumptions. The goal of learning P (Y |X) where X = (X1..., Xi), the Naive Bayes algorithm makes the assumption that each Xi is conditionally independent of each of the other Xk's given Y, and also independent of each subset of the other Xk's given Y . The value of this assumption is that it dramatically simplifies the representation of P (X | Y), and the problem of estimating it from the training data. We can drive the Naive Bayes algorithm, assuming in general that Y is any discrete-valued variable, and the attributes X1... Xn are any discrete or real valued attributes. Our goal is to train a classifier that will output the probability distribution over possible values of Y , for each new instance X that we ask it to classify (Kantardzic, 2003). The expression for the probability that Y will take on its kth possible value, according to Bayes rule, is

$$P(Y=y_k|x_{1\ldots}x_n) = \frac{P(Y=yk)P(X1\ldots\ldots Xn|Y=yk}{\sum P(Y=yj)P(X1\ldots\ldots Xn|Y=yj} \quad \ldots\ldots\ldots\ldots (3.10)$$

Where the sum is taken over all possible values y j of Y. Now, assuming the Xi are conditionally independent given Y, we can use equation (1) to rewrite this as

$$P(Y=y_k|x_{1\ldots}x_n)= \frac{P(Y=yk)\pi iP(X1\ldots\ldots Xn|Y=yk}{\sum P(Y=yj)\pi jP(X1\ldots\ldots Xn|Y=yj} \quad \ldots\ldots\ldots\ldots (3.11)$$

Equation (2) is the fundamental equation for the Naive Bayes classifier. Given a new instance X new = hX1... Xni, this equation shows how to calculate the probability that Y will take on any given value, given the observed attribute values of X new and given the distributions P (Y) and P (Xi| Y) estimated from the training data. If we are interested only in the most probable value of Y , then we have the Naive Bayes classification rule:

$$Y \leftarrow \arg max_{yk} \frac{P(Y=yk)\prod iP(Xi|Y=yk}{\sum jP(Y=yJ)\prod iP(Xi|Y=yj} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.12)$$

$$\mathrm{Yy_k|X_{1\ldots}X_n)=} \frac{\mathrm{P(Y=yk)\prod iP(X1\ldots\ldots Xn|Y=yk}}{\mathrm{\sum P(Y=yj)\prod iP(X1\ldots\ldots Xn|Y=yj}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.13)$$

Which simplifies to the following (because the denominator does not depend on yk).

$$Y \leftarrow \arg\max P(Y = Yk)\prod i\, P(Xi|Y = Yk) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.14)$$

3.9. METHODOLOGY

Methodology is the way that the stated problem that raised in the chapter one of this research solved procedurally. Data mining techniques and architecture is the essential tool in the environment of health for the purpose of prediction and clustering. It is known that data mining techniques contain different types of models to predict and classify the problem of maternal mortality rate that occurs during labor and delivery. There are models which could solve the problem of predicting and creating a model which discussed on the literature review parts of the chapter two sections. Therefore, the researcher should depend on the model which is highly useful for the purpose predicting using the more applicable and recent model of Hybrid model. Hybrid means a thing made by combining two different elements. In this paper, hybrid classification model refers to a combination of two data mining tasks, which are clustering and classification in effort to obtain higher accuracy result (Rozilah et al., 2013).

In Hybrid DM processes model data exploration, understanding the data, cleaning, preprocessing and loading the prepared data into weka version 3.7.5 is the first task and applying the selected algorithm is the road to achieve the objective of this study. The KDP processes are one of the important and necessary one before applying the data mining model (Fayyad et al., 1996). The Hybrid DM processes model is not far from the KDP processes due to the difference between them is not a lot. The development of academic and industrial models has led to the development of hybrid models, i.e., models that combine aspects of both it was developed based on the CRISP-DM model by adopting it to academic research, but Hybrid model a more general than CRISP DM. As a general the hybrid model providing more general,

research-oriented description of the steps and introducing a data mining step instead of the modeling step. Due to this Hybrid is becoming more preferred one (cios et al., 2007). Therefore, Hybrid DM processes model is known by the following descriptive steps and the all applicability of this model would be on the maternity ward data base but ideally understanding the processes like below.

1) Understanding of the problem domain: -This initial step involves working closely with domain experts to define the problem and determine the project goals, identifying key people, and learning about current solutions to the problem. It also involves learning domain-specific terminology. A description of the problem, including its restrictions, is prepared. Finally, project goals are translated into DM goals, and the initial selection of DM tools to be used later in the process is performed.

2) Understanding of the data: -This step includes collecting sample data and deciding which data, including format and size, will be needed. Background knowledge can be used to guide these efforts. Data are checked for completeness, redundancy, missing values, plausibility of attribute values and Noise. Finally, the step includes verification of the usefulness of the data with respect to the DM goals.

3) Preparation of the data. This step concerns deciding which data will be used as input for DM methods in the subsequent step. It involves sampling, running correlation and significance tests, and data cleaning, which includes checking the completeness of data records, removing or correcting for noise and missing values, etc. The cleaned data may be further processed by feature selection and extraction algorithms (to reduce dimensionality), by derivation of new attributes (say, by discretization), and by summarization of data (data granularization). The end results are data that meet the specific input requirements for the DM tools selected in Step 1.

4) Data mining. Here the data miner uses various DM methods to derive knowledge from preprocessed data.

5) Evaluation of the discovered knowledge. Evaluation includes understanding the results, checking whether the discovered knowledge is novel and interesting, interpretation of the results by domain experts, and checking the impact of the discovered knowledge. Only approved models are retained, and the entire process is revisited to identify which alternative

actions could have been taken to improve the results. A list of errors made in the process is prepared.

6) **Use of the discovered knowledge.** This final step consists of planning where and how to use the discovered knowledge. The application area in the current domain may be extended to other domains. A plan to monitor the implementation of the discovered knowledge is created and the entire project documented. Finally, the discovered knowledge is deployed.

3.10. TOOLS AND TECHNIQUES

3.10.1. WEKA

WEKA is one of the data mining tool which is useful in the processes of finding a precious knowledge in data. WEKA stands for The Waikato Environment for Knowledge Analysis. WEKA project was funded by the New Zealand government from 1993 up until recently. The development of Weka was funded by a grant from the New Zealand Government's Foundation for Research, Science and Technology. WEKA provides many different algorithms for data mining and machine learning. Is open source, freely available and plat form independent. Also, easy to use and flexible. This software is the most downloable since April 2000 (Zdravko, 2005).

WEKA is a collection of classifiers (machine learning algorithm) for data mining tasks. The algorithm can either be applied directly to dataset or called from your own Java code.it contains tools for data pre-processing, classification, regression, clustering, association rules and visualization. A set of data items, the dataset, is a very basic concept of machine learning. A dataset is roughly equivalent to a two-dimensional spreadsheet or database table. In WEKA, it is implemented by the weka. Core. Instances class. A dataset is a collection of examples, each one of class weka. core. Instance. Each Instance consists of a number of attributes, any of which can be nominal (one of a predefined list of values), numeric (a real or integer number) or a string (an arbitrary long list of characters, enclosed in" double quotes"). The weka. Filters package is concerned with classes that transform datasets –by removing or adding attributes, resampling the dataset, removing examples and so on. The weka. Filters package is organized into supervised and unsupervised filtering, both of which are again subdivided into instance and attribute filtering (Mark, 2008).

WEKA filters super vised attributes discretize numeric attributes into nominal ones, based on the class information which means Nominal to Binary encodes all nominal attributes into binary (two-valued) attributes, which can be used to transform the dataset into a purely numeric representation (David, 2013). WEKA has several graphical user interfaces that enable easy access to the underlying functionality. The main graphical user interface is the "Explorer". It has a panel-based interface, where different panels correspond to different data mining tasks. The second panel in the Explorer gives access to WEKA's classification and regression algorithms (Witten, 2005).

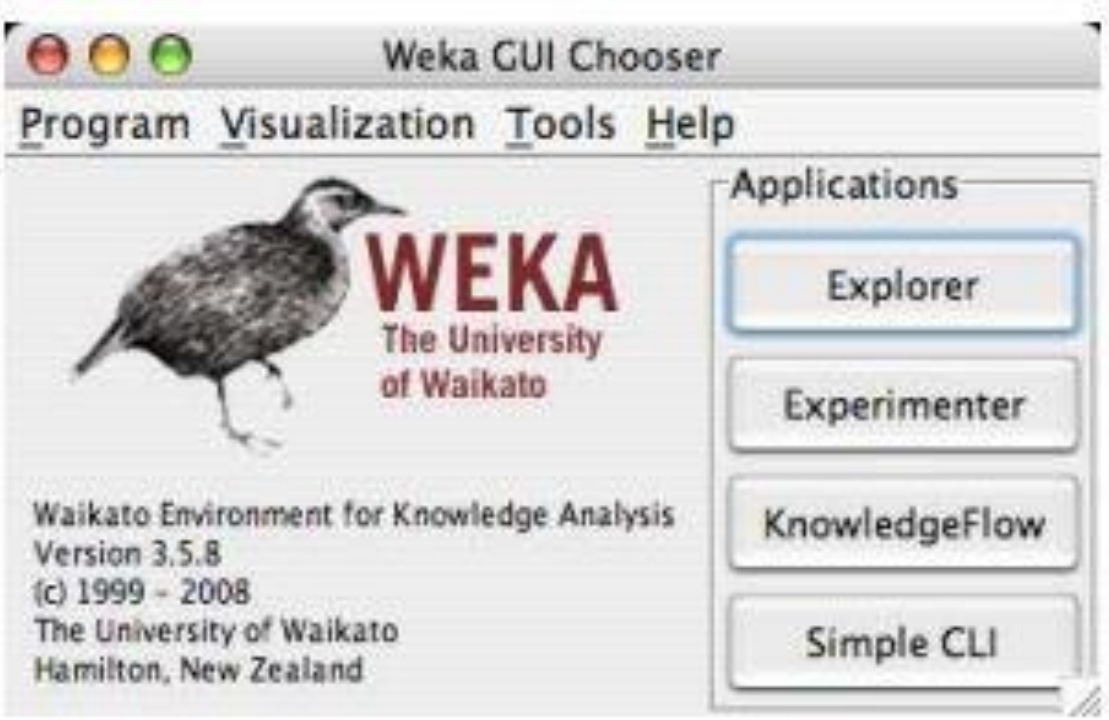


Figure 3.4: The weka GUI interface (Witten, 2005).

Generally, WEKA is a good tool to apply on data set to gain hidden knowledge for one specific purpose. WEKA also accepts the data format by the CSV or comma separated format which is really acceptable format for weka software. Also weka can integrate with different types of programming language which is helpful to apply and use the founded knowledge in the form computer system. The result of data that are inserted into WEKA and developing the interface with visual basic programming is the final achievement of hybrid model.

3.10.2. MICROSOFT VISUAL BASIC.NET TOOL

Visual basic .net is one of the object oriented programme and useful in our todays programme. Visual basic is very easy to use more than other programming language like c^{++} and Java.one of the thing that explains it's easiness is you can drag and drop the interface that the user need and then connecting the code you need.in java or C^{++} you write a code to design the interface or the object you need to shape but this is easy in visual basic.net since you can access the shape that you need from the visual basic library. The Visual Basic language has undergone dramatic changes. Applications and components written in Visual Basic .NET have full access to the .NET framework, an extensive class library that provides system and application services. All applications developed using Visual Basic .NET run within a managed runtime environment, the .NET common language runtime (Grundgeiger, 2002)

Since its introduction in 1991, Microsoft Visual Basic has enjoyed unprecedented success the reason for this success is First, Visual Basic has excelled as a rapid application development (RAD) environment for corporate and commercial applications, Second, Visual Basic offers a programming language and development environment noted for its simplicity and ease of use, making it an extremely attractive choice for those new to programming. Therefore, visual basic is a user friendly programme and also has the ability to connect different databases (Steven.et al., 2001). Lastly taking the most relevant attribute for the prediction of maternal mortality rate and developing the prototype interface which displayed the result for physician is developed by visual basic tool.

CHAPTER FOUR

DATA PREPARATION AND BUSINESS UNDERSTANDING

In real world data is stored and used for different types of purposes in any environment. Data can be stored and saved on different types of computer with different data formats. The world is full of data or like ocean of data definitely. Even though the world is full of data or mass number of data not all data is purposeful. Data should be cleaned and prepared to provide a useful knowledge. Data preparation is more time consuming than data mining. Even though, it is time consuming the prepared data is very vital for the accuracy and performance of the models which is developed for the maternal mortality rate.

Data mining has many methods and techniques to solve the real world problem. Hybrid model is the first choice for this research among the models that is described in data mining. Therefore, it is obligation to make applicable the inner six step processes of it. Hybrid model starts from understanding of the data to use of the discovered knowledge (Cios.et al, 2007). Data preparation is the third step and very important for the next step four, five and six. The other steps is applicable under the next chapter five of this study.

Real world data is impure, noisy, incomplete and inconsistence if we did not prepare it well. Data preparation generates a dataset smaller than the original one, which can significantly improve the efficiency of data mining. Data preparation generates quality data, which leads to quality patterns this can be done through filling the missing value, anomalies, correcting errors, reducing the ambiguity and changing the format into the same format will be the task. It can be understood that data preprocessing, cleaning, and preparation is not a small task (Taylor & Francis, 2003). The purpose of data preprocessing is to clean selected data for better quality. In general, data cleaning means to filter, aggregate, and fill in missing values (imputation). By filtering data, the selected data are examined for outliers and redundancies. Outliers differ greatly from the majority of data, or data that are clearly out of range of the selected data groups (David & Delen, 2008).

Generally, the purpose of this chapter is preparing the quality data of from the maternity ward data base by following hybrid DM processes model i.e. the all steps of hybrid DM processes

model with a data mining techniques. This means discussing the idea of data preparation with maternity ward data which is very crucial steps with in this research.

4.1. BUSINESS UNDERSTANDING

Basically this is the first step of Hybrid model which is selected by the researcher for this study. Business understanding is investigating about the main problem and other additional information with the domain experts. The selected domain experts that can explain and give additional information concerning the maternal mortality rate is Doctors, Nurses, Medical students and Midwife that is roaming in the interface of maternity ward area. This is study is acquitted to understand the main problem of the area with above listed expert groups and drive the study. A description of the problem, including its restrictions also part of it. Finally, the study goals are translated into DM goals, and the initial selection of DM tools to be used later in the process is performed according to the researcher's procedure.

According to David and Dursen, (2008) the key element of a data mining study is knowing what the study is for which means with the concept of predicting maternal mortality is only including pregnant woman's which come for the purpose of labor and delivery in the environment of maternity ward. Therefore, pregnant women's which comes for this service is should be understood by the researcher in case of business understanding.

4.2. DATA UNDERSTANDING

Data understanding the initial and important one to get a good result, solve problems with maternal mortality wich can be done through discussing with domain experts and using knowledge of data mining. Understanding the data or what the data looks like or how much the data is useful for the purpose of prediction maternal mortality rate must analyzed under data understanding. Mohammed & ZakiWagner (2013) explains that Data can often be represented or abstracted as an n × d data matrix, with n rows and d columns, where rows correspond to entities in the dataset, and columns represent attributes or properties of interest. Each row in the data matrix records the observed attribute values for a given entity.

Maternity ward data base has structured with row and columns which held only pregnant woman's information. The maternity ward data base has huge number of records which is in the form of matrix means rows and columns. The data has 23 attributes (card no, hospital name, date of referred, s.no ,Patient name, address ,card noage ,diagnosis ,date of admission ,date of

discharge, length of day ,condition on discharge mode of delivery, indication ,date of delivery ,alive/dead, sex, weight ,PICT, APGAR score, Ballard score, Referral/not, Postpartum, FP use ,not Induction ,Augmentation , Remark, Trombocytopenia) and 5628 rows which only includes woman's information. But fourteen attributes are selected with the value by discussing with the domain experts. Accordingly, fourteen independent variable are selected as the Table 4.1.

Table 4.1: Selected Attributes

NO	Attribute@relation maternity ward data	
1	Address	Inner city, Debub, Rural Area and other
2	Indication on procedure	Normal labor, Poor maternal and Brady Cardia
3	Out come on delivery	Alive and Dead
4	Baby sex	M and F
5	PICT	NR,R
6	APGAR score	Numeric value
7	Mothers BP	High BP and Normal BP
8	Mother's Age	Numeric value
9	Mode of delivery	SVD, SVED and Abortion.
10	Diagnosis	Before and After Delivery
11	Length of stay	Numeric value
12	Condition on discharge	Good and not good
13	Trombocytopenia	Normal and medium
14	Baby weight	Numeric value

4.2.1. DATA COLLECTION

As indicated in section 4.2, it is a bare fact that the concept of data mining doesn't exist without data. For this purpose the researcher used a letter of recognition from the head department of Information Science which makes the researcher more official so that the researcher can contact the domain expert of the maternity wards. Data is important part in data mining since data holds the main hidden information for the study. The data that is used for this research purpose is a secondary data which is found in the maternity ward and encoded by the worker of the

maternity ward daily. The data is stored on the data base in the form of maternity ward data. This data is starting to store from the year that the maternity ward starts to work with Jimma University specialized hospital. The maternity ward and the goals are discussed on the chapter two of this study which is highly focused on the life of pregnant women's. Also this organization is works without any payment and store data for the purpose of pregnant women's information during labor and delivery. According to the domain expert (Dr Dajane) that the researcher contacts the data is stored for more than three years (2014-2017). The data set is 5628 with 23 attributes which only stores pregnancy women's information.

As part of the curse of dimensionality, there were lots of missing values under some of the columns and there were also data items such as different values which is far from the concept of the data, data values in date and time format which don't serve the purpose of mining. Moreover, there have been attributes which are not clear for what they stand for and which the person who gave the data couldn't explain what they are for. Moreover, there were attributes containing redundant values such as birth date, birth date in Ethiopian Calendar, Age in Months, Age in Years and double written "&" sign.

Therefore, describing each of them at this point might not be necessary as there were many of them excluded as part of the tedious process the researcher followed to filter out the more relevant attributes that are important for analysis with the help of domain experts in the area and literatures done on related areas.

4.2.2. ATTRIBUTE SELECTION

With the concept of data mining the word attribute selection is not easy since the performance and accuracy of the model depends on. Attributes selection is the most important stage for the model that is going to developed for the purpose of prediction. This is not an easy step it needs a lot of domain experts discussion beside data. The attribute selection can be done by Weka but what if the Weka did not select the attribute that is not supported by the domain expert means the attribute which the is selected manually before inserting into Weka, therefore discussing and analyzing those attributes which only the researcher manually selected is important.

Most machine learning algorithms are designed to learn is most appropriate attributes for making their decisions i.e Part, J48 and Naïve Byes. Naive Bayes assumes by design that all attributes are conditionally independent of one another, an assumption that is just right for

random “distracter” attributes. Naïve Bayes pays a heavy price in other ways because its operation is damaged by adding redundant attributes. The other J_{48} algorithm inevitably reach depths at which only a small amount of data is available for attribute selection and the dataset were bigger it wouldn’t necessarily help you’d probably just go deeper. Because of the negative effect of irrelevant attributes on most of machine learning schemes, it is common to precede learning with an attribute selection stage that strives to eliminate all but the most relevant attributes. The best way to select relevant attributes is manually, based on a deep understanding of the learning problem and what the attributes actually mean (Frank.et al., 2011).

Generally selecting the attribute selection manually is a good one for the validation. Out of the 23 attributes(card no, hospital name, date of referred, s.no ,Patient name, address ,card no age ,diagnosis ,date of admission ,date of discharge, length of day ,condition on discharge mode of delivery, indication ,date of delivery ,alive/dead, sex, weight ,PICT, APGAR score, Ballard score, Referral/not, Postpartum, FP use ,not Induction ,Augmentation ,Remark, Trombocytopenia) of the original data set, 14 attributes (including the class attribute) which are believed by the domain experts to have significant contribution in predicting the maternal mortality rate, which is the focus of this research, have been selected. Additionally the researchers completely deleted un wanted attribute which is believed by the domain experts un necessary. Those are the list of attribute that the researchers deleted manually out of 23 attributes, ” card no”, “date of admission”, “date of discharge”, “date of delivery” are un selected attribute even though they found on the “maternity ward data base”.

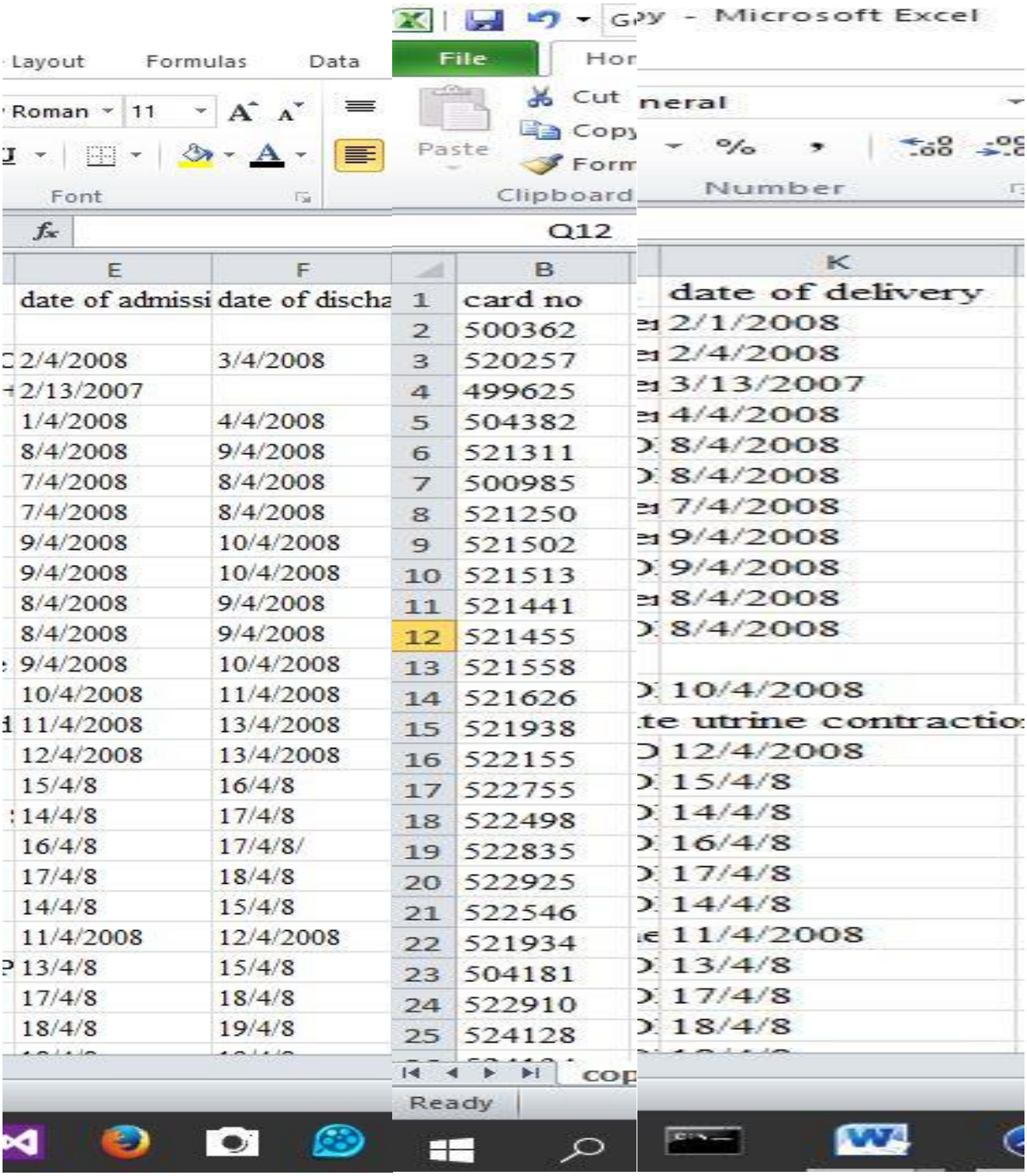


Figure 4.1: Deleted attribute from the data

Figure 4.2: prepared data on the notepad

4.2.3. INSTANCE SELECTION

Not all instances are full or accurate from the maternity ward data base. This could be because of the data encoder. The data encoder is anyone who is professional with IT, means they don't know any concept about the medical term. They only inserted the information that is filled with health professional, during this there is a miss-understanding of hand writing or un clear information which is only acceptable by the domain experts. The biggest point is that there is a missing value or attribute during data encoding and the effect is very high on this study. The researcher investigated that among the 5628 instances, 4218 instances are valuable according the discussion of the researcher and domain experts. Even from this number building a predictive model requires to give the learner algorithm with a training set that have all instance whose outcome or dependent attribute (class label) is not missing. Instance with missing values for outcome class are not useful for predictive model building in data mining because classification algorithms of data mining learn how instance were classified under the different classes. The classes are not existing means the algorithm learns nothing from these instance. As stated by Han and Kamber (2006), records without class labels (missing or not entered) should be ignored, provided that the data mining task involves classification. For this study 4218 instance can make predictive model whereas the left instance is not selected due to noisy, inconsistency and spelling error problem and beside this the behavior of the algorithm that the researcher uses also not need more than this instances for model validation and model development (Ian et al.,2006).

4.2.4. DATA DESCRIPTION

Data description knowing the data contextually and typically.Data has its own meaning and categorization which to mean Data can be numeric, nominal, integer and real number format. The maternity ward data base has 23 attributes which is numeric, integer and nominal depending on the data saved in the form excel. But the researcher only considers on 14 attributes which the domain expert believes on it and the prediction of maternal mortality rate would be valid. Therefore, the data that is identified for this study is figured like below by using the Table 4.2.

Table 4.2: Data description

No	Attribute	Description	Values	Types
1	Address	The address of pregnancy woman	Rural Area, Other ,Debub and Inner city	Categorical
2	Mothers Age	The age of the pregnant woman	Numeric age value	Numeric
3	Diagnosis	Diagnosis at presentation	Before Deliver and After Delivery	Categorical
4	Length of stay	The duration of stay after the pregnant women delivered	Numeric length of stay value	Numeric
5	Mode of delivery	Mode of delivery	SVD, SVED and Abortion.	Categorical
6	Indication for procedure	Indication for procedure	Normal labor, Poor maternal and Brady Cardia	Categorical
7	Out come on delivery	Out come on delivery	Alive and Dead	Categorical
8	Baby sex	The sex of baby	Female, Male	Categorical
9	Baby weight	The weight of the new born baby	Numeric weight value	Categorical
10	PTC result	HIV test result	NR,R	Categorical
11	APGAR score	The score for the new born baby	Numeric value	Numeric
12	Mothers BP	Mothers blood pressure	High BP, Medium BP and Normal BP	Categorical
13	Thrombocytopenia	The score that show woman's blood platelets	Normal and Medium	Categorical
14	Condition on Discharge	Status of patient before discharge	Good and Not Good	Categorical

4.3. DATA PREPROCESSING

4.3.1. DATA SUMMARIZATION

Descriptive data summarization techniques used to identify the typical properties of your data and highlight which data values should be treated as noise or outliers (Jiawei and Micheline, 2006). Furthermore, missing values can easily be identified through this technique which in turn facilitates the next phases of data preparation.

Descriptive data summarization is evaluating with respect to their mean, median, mode and mid-range for numeric attributes found in the data set. The researcher identified that "age of mother", "baby weight", "length of stay", should be under numeric data value. According the Table 4.3 below the mean, standard deviation, max and min value is illustrated.

Table 4.3: Exploratory Data Analysis

Exploratory Data Analysis				
		Mothers Age	Baby Weight	Length of stay
N	Valid	4218	4218	4218
	Missing	2%	0	0
Mean		25.65	3.371	1.382
Std. Deviation		5.026	1.652	0.697
Minimum		14	1.9	1
Maximum		47	12.2	3
Outliers		<17.3 and 43.3>	<0.5 And 4>	>34 days

4.3.2. DATA CLEANING

Data cleaning, also called data cleansing or scrubbing deals with detecting and removing errors and inconsistencies from data in order to improve the quality of data. The researcher is discuss and clean the data depending on the attributes types. Imagine there are "names" under the attribute of "m-ages" which is miss placed due to typing or basically errors. Additionally there are also there are signs "@" or "&" which the encoder just put instead of living the columns null. Those factors raise a problem for the model validation in the experiment scenario. Due to this reason it is important to clean and replace the un necessary value by discussing with the

domain experts, as an example under the column address 01 kebele and under baby weight column kgs.

When multiple data sources need to be integrated, example in data warehouses, federated database systems or global web-based information systems, the need for data cleaning increases significantly (Kamber, 2006). This is because the sources often contain redundant data in different representations. In order to provide access to accurate and consistent data, consolidation of different data representations and elimination of duplicate information is necessary. Due to this case the maternity data is stored by different class names which are on the same attribute i.e APH, SVD, Eclampsia, Diagnosis and SVED. These are selected into the same title to get a cleaned data by discussing with the domain experts.

The data cleaning tasks performed to raise the data quality to the level required by the selected analysis techniques involves selection of clean subsets of the data and the insertion of suitable defaults. To this end, through discussion has been made with the domain experts and found out missing attribute values at the time of data entry are recorded as "unknown" and those records which the attribute is irrelevant simply left as a blank assuming that it would be obvious. To fix these problems some records with missing or unknown values for significant number of attributes are removed from the dataset. Noisy values for attributes are also deleted and set to blank cause on the data bases there were many encoders name on the raw of the database, like "masereti","&&&","1", which is meaningless and those are removed during data cleaning.

4.6.2.3. HANDLING MISSING VALUES

Missing data is a problem that continues to affect data analysis methods (Larose, 2005). The absence of information is rarely beneficial. All things being equal, more data is almost always better.

According to Larose (2005) missing values may occur for several reasons, such as malfunctioning measurement equipment, lack of consistency with other recorded data and thus deleted, or respondents in a survey may refuse to answer certain questions such as weight, height or religion and data may not be recorded due to misunderstanding. But those missing values needs to be given significant attention.

To deal with missing values, alternatives are suggested by Larose (2005) and Chackrabarti et al., (2009). These are:

- Ignore the missing value
- Replace the missing value manually
- Replace the missing value with a global constant to fill in the missing value
- Replace the missing value with some constant, specified by the analyst
- Replace the missing value with the field mean(for numerical variables) or the mode (for categorical variables)

Table 4.4: Handling missing values Experiment

No	Attribute	Valid	Missing value	Action taken
1	Address	4218	0	No action taken
2	Mothers Age	4241	0	No action taken
3	Diagnosis	4113.77	(127.23)3%	Filled by discussing with domain expert
4	Length of stay	4022.3	(211.7)5%	Filled by discussing with domain expert
5	Mode of delivery	3641.24	(592)14%	Filled by discussing with domain expert
6	Indication for procedure	4191.66	(42.34)1%	Filled by discussing with domain expert
7	Out come on delivery	4218	0	No action taken
8	Baby sex	4218	0	No action taken
9	Baby weight	4191.66	(42.34)1%	Removed
10	PTC result	4234218	0	No action taken
11	APGAR score	4106	(127.02)3%	Filled by discussing with domain expert
12	Mothers BP	4218	0	No action taken
13	Thrombocytopenia	4218	0	No action taken
14	Ballard score	3641.24	(592) 14%	Filled by discussing with domain expert

4.6.2.4. RESOLVING INCONSISTENCIES

The two possible causes for the discrepancies detected in the fields of selected attributes are human error in data entry and the design of the values of attributes of the database with no predefined values. The problem associated with existence of inconsistencies is that they reduce the quality of the final model and makes learning difficult for the algorithms (Han and Kamber, 2006).

Discrepancies were detected while extracting statistical summaries of attribute values. There are invalid values entered in the database. For instance under the field 'Adress', the terms "Jimma", "jim", "JIMMAA", "JImmaa" etc. were used to describe for the address of the woman's who came from "Jimma". Therefore for the sake of consistency, the researcher corrected them as one category "Jimma". There are also other cases of expressing another addresses which has the same behavior under the column of mothers address .For instance addresses like 'jimma','limu' 'kosa','debub','shabe','sombo','kersa','limmu','kersa','chorabotor','kesa','shebe','O/nada','s/ sonbo','dedo','SOKORU','Gomma','l/seka','mana',T/afeta,'Kersa','sigmo'.

Therefore, the researcher understands those listed address could be due to error while encoding and corrected with "Jimma", "Limu", "Kersa", "Chirobotor", "Shabe", "OmoNada", "ShabeSombo", "Dedo", "Sokoru", "Mana which is the same with Gomma", " Limmu Seka","Tiro Afeta","Sigmo". Additionally the under the attribute of mode of delivery there are words which almost the same i.e 'SVeD','SVD','Svd','Svd',' SVED' is corrected with word "SVD". VD + Eclampsia or Eclampsia + VD is almost the same, therefore this is almost corrected into Eclampsia + VD. On the top of this under the column "APGAR score" there are data which is written in the form of this "7&&8","7&11", this is just corrected into "7&8" cause the "&" sign which is written two times is the encoders error and "7&11" is changed to "7&10" cause the APGAR score of the baby never above the 10 according to the domain experts suggestion.

Generally, the researcher categorized the above listed class attribute into "Inner City","Debub" ,Rural Area and "other" for the Address attribute, this is done by discussing with domain expert since the distance which the pregnant woman comes has effect on the predicting maternal mortality rate. And for the "Diagnosis" attribute the researcher make into "before Delivery" and "after Deliver" instance for the column of Diagnosis instances this is again done by discussing

with the physician which diagnosis found before and After the pregnant woman would affect her life.

4.2.6.5. HANDLING OUTLIERS

A database may contain data objects that do not comply with the general behavior or model of the data. These objects are considered as outliers. Deviation - based methods identify outliers by examining differences in the main characteristics of objects in a group. The degree to which numeric data tend to spread is called the dispersion, or variance of the data. The most common measures of data dispersion are range, based on quartiles, the inter quartile range and the standard deviation. Box plots can be plotted based on the five number summaries and are a useful tool for identifying outliers (Han and kamber, 2006).

Accordingly, the outlier values within the attribute of the dataset used for the research especially for numeric type attributes were explored and approached based on recommendations from different data mining literatures to handle the outlier values. Detecting the outliers is important cause they minimize the accuracy of the model and outliers detection can be categorized under the gross error from the type of errors. Outliers are data measurements occurring from gross errors which to mean weather the recorded data is useful or not (Analysis of Errors, 2013).

As stated in Han and kamber (2006) a common rule of thumb for identifying suspected outliers is to single out values falling at least 1.5* IQR above the third quartiles or below the first quartile. In other words it is to mean that the values outside the limits: Q3+ (1.5*IQR) and Q1-(1.5*IQR) will be considered outliers values. Based on the recommendations, Baby weight, Mothers Age and length of stay which are numeric, had outlier as stated in the Table 4.3.

4.2.8. DISCRETIZATION

According to Chackrabarti et al. (2009), data discretization techniques are used to reduce the number of values for a given continuous attribute by dividing the range of the attribute into intervals. This leads to a concise, easy-to-use, knowledge-level representation of mining results. Among the available data discretization techniques, binning has been selected to discretize the numeric attributes in to ranges of values for the independent variables and in to classes for the dependent variables.

Binning is a top-down splitting technique based on a specified number of bins (Chackrabarti, et al., 2009). The attribute values can be discretized by applying equal-width or equal-frequency binning which is done by WEKA tool.

4.2.8.1. DISCRETIZING THE VALUES OF MOTHERS AGE ATTRIBUTE

In the chapter one of this study that women's never get pregnancy under the age of thirteen and fourteen, due to this discretizing the attribute of mother's age is important. In order to reduce and simplify the original data, the researcher replaced numerous values of a continuous attribute by a smaller number of interval labels there by reduces and simplifies the original data. This leads to a concise, easy to use, knowledge-level representation of mining result. Concept hierarchies can be used to reduce the data by collecting and replacing low-level concepts (such as numerical values for the attribute age) with high level concepts for instance, adolescent ,adult and old (Han and Kamber, 2006).

Binning is used to scale the values of "Age" attribute. Discretizing the values of mother's age is done by grouping the age value in to eight distinct i.e, "17.3-20.6", "20.6-23.9", "23.9-27.2", "27.2-30.5", "30.5-33.8", "33.8-37.1", "37.1-40.4" and "40.4-43.7". By taking this in to consideration the researcher decided to take the value below under 17.3 years and 43.7 as outliers. The discretized values for the age is stated as Table 4.5

Table 4.5: The discretized age attribute.

No	Age attribute in range	Transformed value
1	"17.3-20.6"	A1
2	"20.6-23.9"	A2
3	"23.9-27.2"	A3
4	"27.2-30.5"	A4
5	"30.5-33.8"	A5
6	"33.8-37.1"	A6
7	"37.1-40.4"	A7
8	"40.4-43.7"	A8

Therefore, 43.7 years is the upper limit for outliers so that age values beyond 43.7 years were treated as outliers. For the lower limit is 25-(1.5*IQR) which is 28-16.5=17.3 So Age values bellow 17.3 years can be considered outliers in this dataset. These grouped six distinct values are used to see patterns in "Age" if any. Therefore, the actual low level values are manually replaced by the corresponding high level grouped values.

4.2.8.2. DISCRETIZING THE VALUES OF BABY WEIGHT

As Chackrabarti, et al., (2009) binning is a top-down splitting technique based on a specified number of bins. The attribute values can be discretized by applying equal-width or equal-frequency binning. The researcher understands the attribute of baby weight is vary depending on records under baby weight column. The registered number for the baby weight is between 1.5kg up to 4.0 kg which means baby weight can be between 1.5 and 4.0 kg. Again by discussing the midwifery and doctors of the maternity ward there is no baby weight less than 1.5 and above 4.0 kg which means any number below or above this number is error or outliers for this attributes. As an example while the researcher is looking into the data there are baby weight which is registered in the form of 0.5 mg which is outliers. Therefore the researcher decided to bin the baby weight categorized into five distinct values to form a total of five bins "1.5-2.0","2.1-2.5","2.6-3.0","3.1-3.5","3.6-4.0".The rest data value between 1.5 and 4.0 are equally portioned by the difference of five

Table 4.6: Baby Weight discretization

No	Baby weight attributes in range	Transformed value
1	"1.5-2.0"	W1
2	"2.1-2.5"	W2
3	"2.6-3.0"	W3
4	"3.1-3.5"	W4
5	"3.6-4.0"	W5

4.2.8.3. DISCRETIZING THE ATTRIBUTE OF BABY SEX

The data is full of noisy still, due to there is still data which is encoded wrongly under the column of baby sex attribute. The baby sex column is written “M & M” which is both of them are male and “F&F” which is both of them are female, “F&M” and vice versa, this double writing is the same under attribute weight. Therefore splitting these rows by making copy paste into the next row and make value of baby sex and baby weight into different place is done.

4.2.9. DATA TRANSFORMATION

According to Ian et al. (2006), data transformation involves transforming or consolidating the data to a form appropriate for mining. Data transformation usually involves data smoothing, generalization of data, normalization of data, aggregation of data, and attribute construction.

In this study, generalization of data was used in order to give a good accuracy for the class attribute there are instances which is need transformation under the column of “length of stay” and “baby weigh” under this two column there are instance which is written in the form of Hours instead of writing in a day i.e 12 hrs,13 hrs. But by the domain expert said by default once the woman is delivered the should stay even for one day and under the column baby weight the baby is written in form of “Kg” which is for weight but there are still some instances which is written in the form of “gm.” which is less than “Kg”. As a general the researcher calibrated according to the Table 4.7.

Table 4.7: Class Attributes Transformation.

No	Attribute of	Transformed value
1	“gm”	Removed
2	“12hrs”,”13hrs”	Converted to one day
3	Jimma city, Bosa kitto, shabe,sombo	Inner city
4	Seka, Agaro, Omo, Nada, Gatira, mana, Sokoru, dedo, limmu Genati	Rural Area
5	Tercha,Saja,Telo,Sigmo,limmu kossa	Debub
6	Mettu,Beddele,wollege,waliso	Other
7	120:80	Normal BP
8	140:170	High BP
9	124:90	Medium BP

The above Table 4.7 shows the class which is transformed into other value. From the table starting from number 3-6 is the transformation which is made for the Address. Number two the instance which is converted into one day under the attribute of ”Length of stay”. The instant from no 7-9 is the Blood pressure information for the pregnant woman’s according to WHO standards.

4.2.10. DATASET FORMAT

Weka needs data to be prepared in some formats and file types. The data sets provided to this software are prepared in a format that is acceptable for Weka software. Weka accepts records whose attribute values are separated by commas and saved in an ARFF (Attribute-Relation File Format) file format which the file is initially from the “Jimma University specialized Hospital Maternity Ward data base”.

In order to prepare the data in such format the records from the Microsoft excel database are saved as a Comma Delimited (CSV) file. Once all processing is completed and the file is converted to .csv format, WEKA either process the .csv format itself or a file in the form of Attribute Relation File Format (arff). For this study the data is given to the software in .arff format for the experiment.

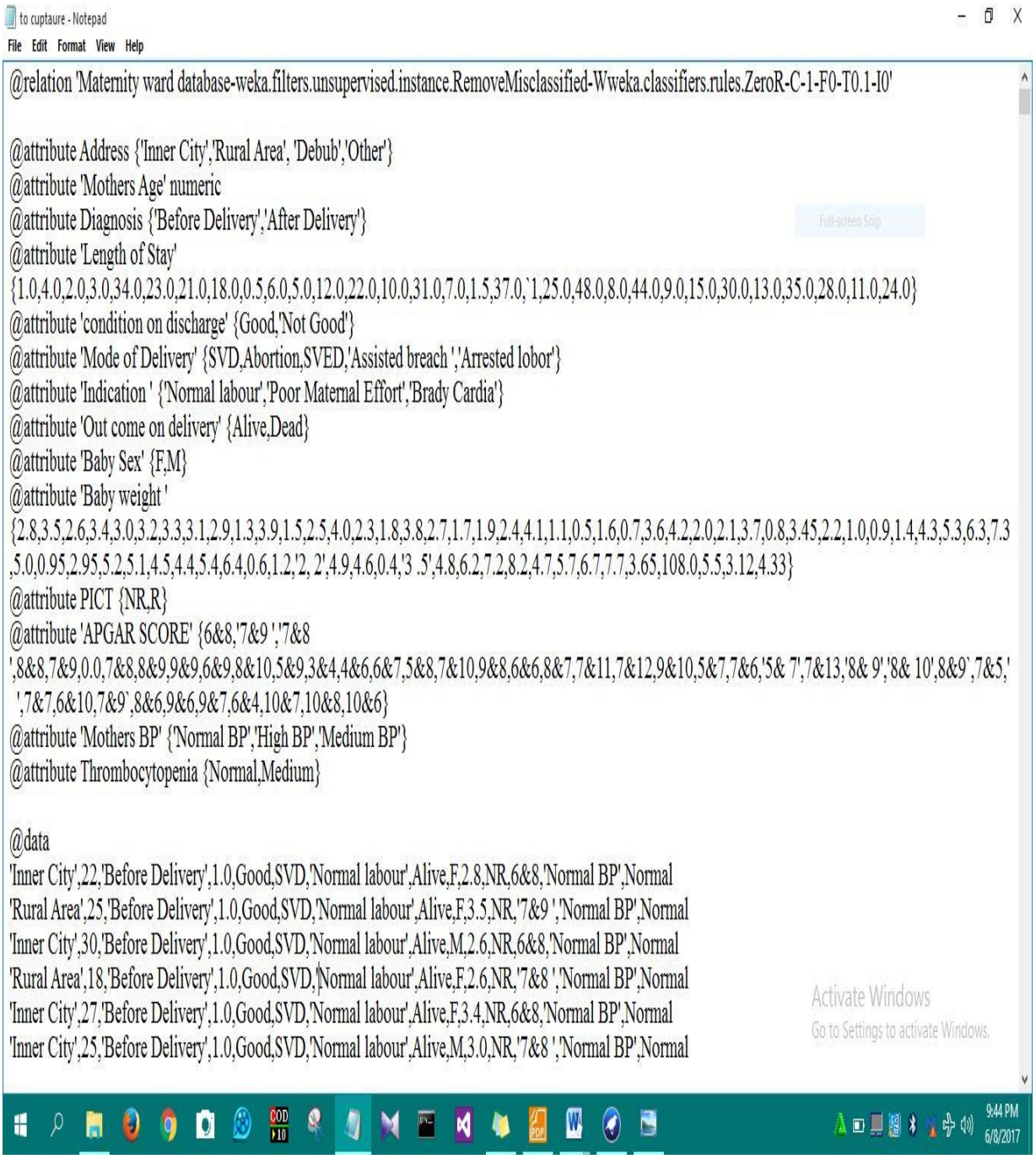


Figure 4.3: The arff data for weka interface

CHAPTER FIVE

EXPERMENTATION

In this study an attempt was made to design a model that enables to predict the maternal mortality by using data mining techniques. This is the final step that the hybrid model is applicable according the procedure of the experiment. Data is already prepared on the preprocessing stage. This data contained the cleaned and error free data is tested in this study to model the prediction task. Experimentation is the interface that tests the data according to the researchers selected algorithm under methodology. The PART, J48 (known C4.5) and Naïve Bayes is the selected algorithm to develop a predicting model under this section. For creating a prediction model on a data set which contains 4218 is prepared for testing and training. For this experimentation task the researcher is discussed the main attribute which could predict the maternal mortality rate.

5.2. ATTRIBUTE ORDERING

Since attribute selection is important in decision tree models, the researcher tried to rank the attribute based on information gain. It was calculated based on entropy value of the attribute. As Witten and Frank (2005) explained information gain is calculated from sum of entropy for every attribute. The formula for calculating intermediate values is:

$$\text{Info}(D) = -\sum_{i=1}^{m} P_i \log_2 P_i \quad \ldots\ldots\ldots\ldots\ldots (5.1)$$

Where, Pi is the probability that an arbitrary tuples in sets of training data D belongs to certain class. Info (D) is also known as the entropy of D. After calculating information gain for each attribute, select the one with the highest information gain as the root node, and continue the calculation recursively until the data is completely classified by J48 algorithms in this case. For the purpose of this research, WEKA is used to compute the information gain and rank according to the importance of the attribute for the classifier. As a result of this, the following figure 5.2 depict ranked order of attribute based on their relevance for the reason that such attributes are very important for later experimentations by excluding the least relevant attributes.

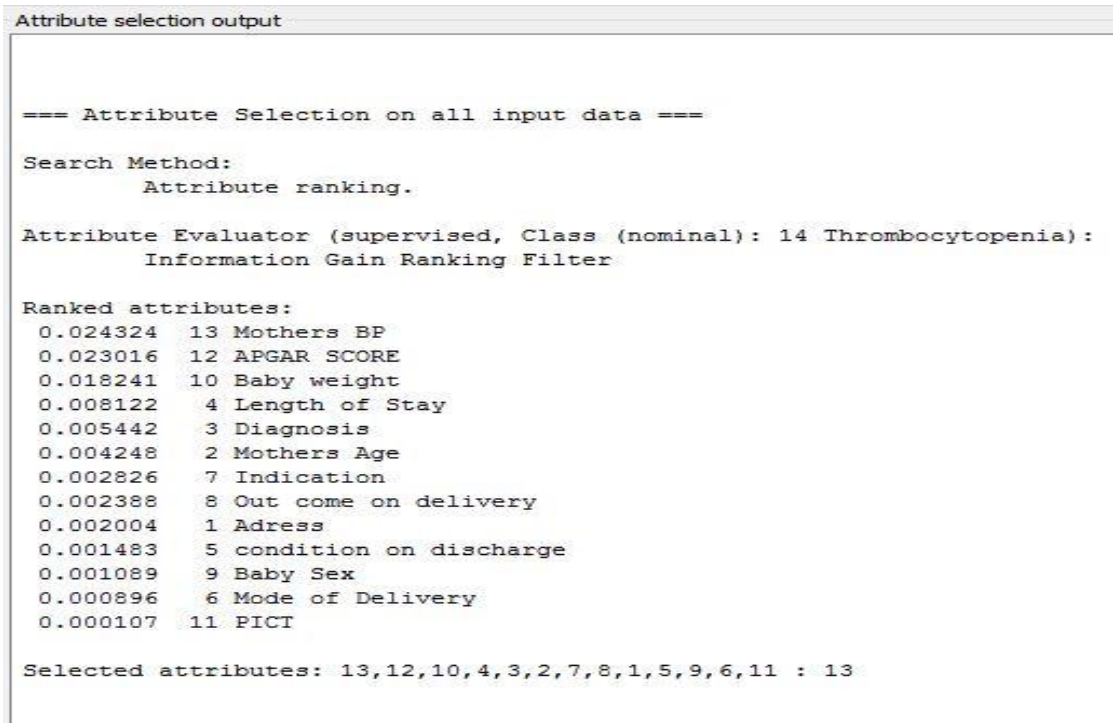


Figure 5.1: The ranked Attributes

Shown in the result of attribute selection using entropy based information gain method of WEKA, out of 14 attributes, the top 10 the most relevant attributes are Mothers BP , APGAR Score, Baby Weight, Length of stay, Diagnosis, Mothers Age, Indication, Out Come on Delivery, Address and condition on discharge .

5.3. EXPERIMENTAL DESIGN

In this study, all experiments are done by using the data which is prepared on the previous stage which contains 14 attributes (Address, Mothers BP, Age, Outcome on delivery, Diagnosis, Baby sex, Baby weight, Trombocytopenia, Length of stay, Indication, PTC, APGAR Score, and Baby sex) and 4218 instances to build a predictive model for the purpose of maternal mortality rate prediction by using hybrid model. Some prerequisite condition such as data cleaning, data loading and attribute selection is held in the previous task and now the left step is applying the experiment part.

The algorithms (PART, J48 and Naïve Bayes) used during both predictive model building and experimentations are found in Weka 3.7.5 version. This version works on many file formats than its antecedents and it is compatible with CSV file format. Thus, no additional effort was

exerted to change the dataset from excel to ".arff" file format which is necessary in the previous versions. The prepared dataset is saved using CSV file extension format. Weka has four testing option which is use training set, percentage split, supplied test set and cross validation set (Richard and Frank, 2004). The Experimental design is applying on those selected algorithms in the data set by using K-fold cross validation i.e 10 fold cross validation for each algorithms. For this study building a classification model for predicting maternal mortality rate is accomplished under these experiments. Beside this evaluating the performance and accuracy of the algorithms to deciding the most performing model for the prediction of maternal mortality rate would be completed for the prototype development.

For the classification purpose and building model the researcher used two methods for testing and training the data set. The first method is percentage split method, where 70 % of the data used as training and the remaining 30% testing. The second method is K-fold cross validation methods, the data was divided into 10 folds for testing and the remaining folds are used as training. Then according the Hybrid model validation and changing the output of the Weka into a computer system which is lastly displaying by using Microsoft Visual basic.

Shortly the experiment is designed to do like the above scenarios. Scenario one works by J48 algorithm by with pruning and un pruning and also for the PART and Naïve Bayes algorithm to find the better accuracy. Under this scenario discussing the accuracy by mean absolute error, accuracy of the model, sensitivity (TPR), Specificity, False Positive Rate, F-measure and area under the ROC curve. The listed performance criteria's are selected by some previous researcher (Mulugeta.et al., 2013).The data which is prepared for the experiment is like the Figure 5.2.

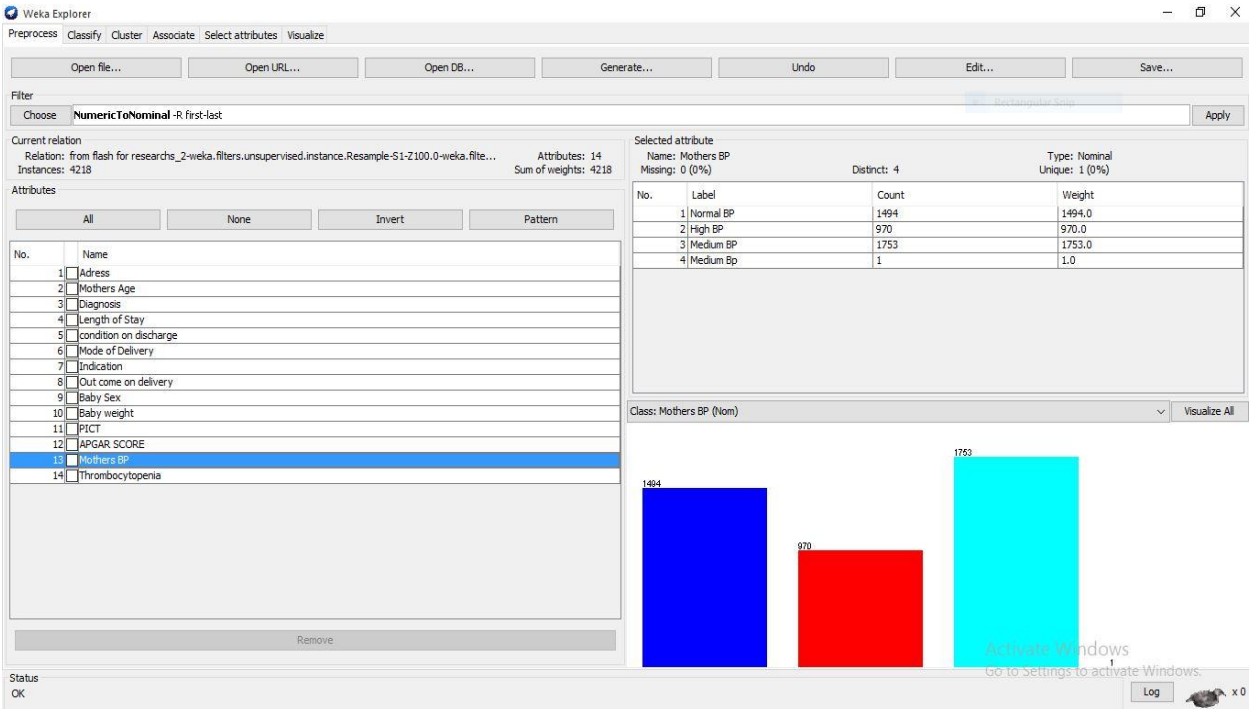


Figure 5.2: Weka Explorer window showing the number of attributes and instances

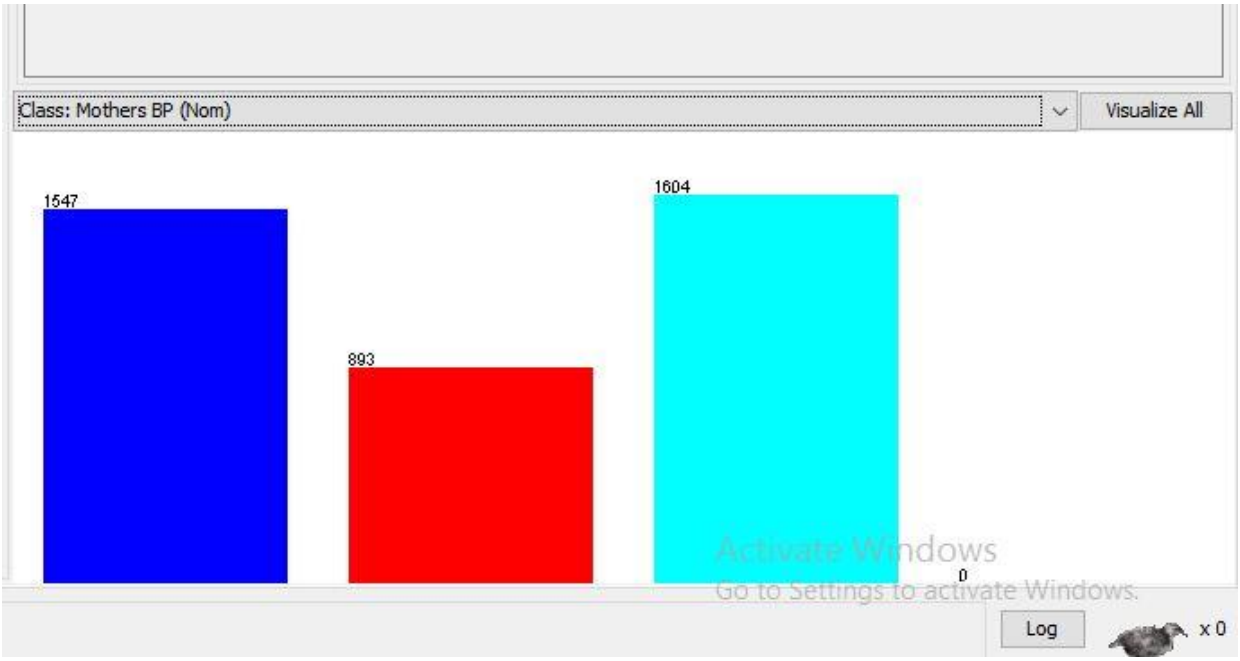


Figure 5.3: Weka Explorer window showing the number of attributes and instances after Resample the data.

The resample techniques help to balance the data so that the data is equally accessed by the algorithm. Resample is in Weka version 3.7.5 and SMOTE techniques are in weka version below 3.7.5 but ideally the same.

5.4. J48 DECISION TREE MODEL BUILDING

The advantage of using Decision Trees in classifying the data is that they are simple to understand and interpret. Decision tree uses the divide and conquer method means they tend to perform well if a few highly relevant attributes exist (Ahmad et al., 2013). Decision tree is very useful for classifying an unknown sample and testing the attribute values of the sample against the decision tree. If then rule is used in these criteria by dividing the valuable knowledge into right and left sides. Also decision tree classifiers known by C4.5 decision tree algorithm but it also known by J48 on the weka environment (Mustapha.et al., 2013).

5.4.1. EXPERIMENT WITH J48 DECISION TREE MODEL BUILDING

According the scenario of under the experiment design dividing the prepared data set into 70% training and 30% for testing then accordingly running the experiment is the task. This means that by deciding the number of iteration that the algorithm would iterate. Therefore dividing the data into two and making the number of cross fold validation 10 is running under this section. According to the above experiment design discussing issues wich would prove the maternal predicting case would be identified by checking the output of the classifiers. This experiment would conduct on the data set of maternity Ward Database with 4218 instances. With respect to this experiment Hybrid model is running beside of this experiment cause in the steps of hybrid model, model creation and evaluation should be conducted under here regarding with picking out the best performing model among the algorithms. Picking out the best attribute which predicts the maternal mortality rate is identified under here which means J48 by switching the parameter with pruning to TRUE and FALSE to form two separate experimental settings under this scenario.

Accordingly the researcher experiment the first J48 algorithm running on the data set. In this experiment, the performance of J48 classifier in predicting maternal mortality rate is evaluated. Two models of the J48 algorithm (pruning and un pruning) models were built using J48 decisions tree with default parameters and all (14) attributes.

5.4.2. EXPERIMENT 1

This is the first experiment of this study which is applied on the data set of 4218 instance according to the above experiment design and the result of the experiment is discussed under the Table 5.1.

Table 5.1: J48 Decision tree result with pruned

Model	J48 Decision tree algorithm pruned
Instances in number	4218
No of leaves	1
Time (sec)	0.14 sec
Average Precision	0.958
Average ROC area	0.538
Average TP rate	0.956
Average TN rate	0.951
Average FP rate	0.951
Accuracy	95.637%
Correctly Classified	4034
Incorrectly Classified	184
Method	70/30 methods

Table 5.2:J48 decision tree with un pruned

Model	J48 Decision tree algorithm un pruned
Instances in number	4218
No of leaves	1848
Time (sec)	0.02 sec
Average Precision	0.987
Average ROC area	0.988
Average TP rate	0.987
Average TN rate	0.226
Average FP rate	0.226
Accuracy	98.7435%
Correctly Classified	4165
Incorrectly Classified	53
Size of the tree	1947
Method	10 fold cross validation

From the above Table 5.1 and Table 5.2 the result of the experiment was represented according to the output of the weka. The above experiment is done on the instances of the 4218 data set and the result of the J48 algorithm is represented too. From the above Table 5.1 and 5.2 the J48 pruned and un pruned decision tree is displayed on the weka to show the predicting maternal cases. The J48 algorithm with pruned decision tree scored the accuracy of 95.67% with the time elapse of 0.14 sec. This means the accuracy of the model relating with classifying the data set of 4218 is approximately error with 4.362 %. In this experiment from 4218 data set 4034 data is classified correctly and 184 data is incorrectly classified from the same data set. These experiment has the ROC area is 0.538 wich is almost greater than five and good to determine the models performance. The confusion matrix of the J48 pruned model looks like below wich looks like in the form of a and b matrix.

Table 5.3: The confusion matrix of the J48 pruned decision tree

=== Confusion Matrix ===		
A	B	<-- classified as
4033	0	a=normal
184	1	b=medium

From the confusion matrix, 4033 records classified as normal from the data set of 4218 and one record were classified as medium. The classifiers also classifies 184 records as normal and zero records as medium

The second experiment is which is done with the un pruned decision tree (True).This experiment again done on the data set of 4218 with the method of 10 fold cross validation with the 70/30. This is the experiment that the researcher founds a most promising result for the purpose of maternal mortality rate prediction of pregnant woman. In this the accuracy of the model is much better than the pruned decision tree classifiers or the first experiment. This model takes 0.02 seconds to build a classifiers model for the predicting purpose. From the 4218 instances 4165 correctly classified and 53 is incorrectly classified. The accuracy of the model is 98.7435 % correctly classified and 1.2565% incorrectly classified. Additionally, on this un pruning experiment of the decision tree 1848 leaves of the tree and 1917 size of the tree. The kappa statics of this model is 0.8334 and the ROC area is 0.988.The ROC area helps to

determine the accuracy and the performance of the model. Again the more the wide area under ROC curve the more the model is accurate.at this stage again a there is a confusion matrix occurrence wich is helps to determine the how well the classifier is. The confusion matrix of the J48 un pruned decision tree looks like below.

Table 5.4: The confusion matrix of J48 decision tree with un pruned

=== Confusion Matrix ===		
A	B	<-- classified as
4026	10	a=normal
43	139	b=medium

From the above Table 5.4 the confusion matrix is used for how well the classifiers are good in the prediction of the maternal mortality rate. This shows form the data means 4218 data set 4026 records were classified as Normal and 139 records were classified as Medium. The classifier also classified 10 records as Medium and 43 records as normal. The above Table 5.4 can be generated like the Table 5.5.

Table 5.5: The confusion matrix of J48.

	Predicting maternal mortality rate		
	Normal	Medium	Total
Maternal status as normal	4026	10	4036
Maternal status as medium	43	139	182
Total	4089	149	4218

Furthermore, evaluating the model based on sensitivity and specificity are very significance for decision making. For that reason, the result of the above confusion matrix indicate that the sensitivity of this test was (4026/4036) = 99.76 % and the specificity was (139/182) = 76.4%. The test indicates that the models appear to be pretty good. Because, based on the evaluation criteria, the classifier correctly classifies clients as Normal status rate with 99.76% accuracy.

To sum up the above detailed confusion matrix of un pruned J48 from the data set of 4036, 4026 or 99.7 % correctly classified as Normal status and 10 or 0.002 % are miss classified. Also from the 182 medium status 139 or 74% are classified as medium status and 43 or 23.6% miss classified.

To conclude the above Table 5.5 for the experiment of J48, the J48 un pruned experiment has a more accuracy and covers a wide area of ROC curve which is very essential to discuss the performance of the model. Again the J48 un pruned algorithm has been scored the accuracy of 98.7435 whereas the pruned J48 algorithm scored an accuracy of 95.6377. Beside this, by classifying a number of instances also J48 un pruned decision tree is better. From this the researcher identified that the J48 decision tree algorithm is a good predictor model more than pruned decision tree algorithm. As a general in this experiment the selected 14 attribute is running with 4218 instances to see which attribute has a factors on the prediction of maternal mortality rate which is raised under the chapter one of this study as a research question.

5.5. EXPERMENT WITH PART RULE INDUCTION MODEL BUILDING

The second data mining classification technique applied in this research was PART Rule induction algorithms. As mentioned in literature review, there are many rule induction algorithms but the researcher selected PART for the reason that PART has the ability and potential to produce accurate and easily interpretable patterns or rules that helps to achieve the research objectives. PART is a separate-and-conquer rule learner like that of decision tree and proposed by Witten and Frank (2005).

To build a rule model induction again a weka contains a set of 4218 datasets which is prepared before. This experiment still drives according the first experiment of J48 algorithm. Accordingly, the procedure of the Hybrid DM processes model still applicable here means discussing the accuracy, TP, ROC and number of rules produced. This helps to decide the best rules for the prediction of maternal mortality depending on the rules that weka produced with the PART rule induction tree. The methodology which is applied first is still applied here which means 10 fold cross validation and (70/30) division of data set. The 10 fold cross validation is dividing the data into ten iteration places which is helpful for the experiment. The parameters are partially adjusted and the 14 attributes are selected again.

5.5.1. EXPERMENT 1

This scenario or experiment is taken on the data set of 4218 by the method of (70/30) for training and testing procedurally. Again in this experiment using the 10 fold cross validation as a second method. Experimenting with the PART tree model is to identify how the model is accurate. The result of the experiment is written according to the table below.

Table 5.6: The PART tree rule induction result

Model	PART tree algorithm un pruned	
	70/30 methods	By 10 fold cross validation
Instances in number	1265	4218
No of rules	105	105
Time (sec)	0.14sec	0.12sec
Average Precision	0.942	0.948
Average ROC area	0.793	0.786
Average TP rate	0.957	0.957
Average TN rate	0.816	0.705
Average FP rate	0.816	0.705
Average recall	0.957	0.957
Accuracy	95.65	95.7
Correctly classified	1210	4037
Incorrectly classified	55	181

Two methods applied on this experiment to measure the best performing model for the PART tree rule induction method. On this scenario the models scored the accuracy of 95.65 % on the 70/30 methods. On this method from the 1265 instances 1210 correctly classified and 55 instances incorrectly classified. In this method instances take 0.14 sec takes to classified .On the second method means by making the 10 fold cross validation the model classified instance correctly 95.7% from the 4218 instances. For the second method time taken to classified the instances are 0.12 sec and 4037 classified correctly and 181 in correctly classified. With time and accuracy the un pruned rule induction part tree is way better than the 10 fold cross validation.

Table 5.7: The confusion matrix of PART tree for un pruned

=== Confusion Matrix ===		
A	B	<-- classified as
3989	47	a=normal
134	48	b=medium

The above confusion matrix or Table 5.7 shows that from the dataset of 4218 instances 3989 records are correctly classified as normal and 48 records are correctly classified as medium. From this data set again 134 correctly classified as normal and 47 records are correctly classified as medium. From this confusion matrix show how better the un pruned part tree rule induction better than for the selected two methods wich is by 10 fold cross validation. The above confusion matrix can be evaluated as Table 5.8.

Table 5.8: The general confusion matrix of PART rule induction.

	Predicting maternal mortality rate		
	Normal	Medium	Total
Maternal status as normal	3989	47	4036
Maternal status as medium	134	48	182
Total	4123	95	4218

For the above PART rule induction Table 5.8 again evaluating the model based on sensitivity and specificity are very significance for decision making. For that reason the result of the above confusion matrix indicate that the sensitivity of this test was (3989/4036) = 98.86 % and the specificity was (48/182) = 26.37%. The test indicates that the models appear to be somewhat good next to the J48 algorithm. Because, based on the evaluation criteria, the classifier correctly classifies clients as Normal status rate with 98.86% accuracy.

To sum up the above table depending on the maternal status as normal and maternal status as medium out the general data 3989 correctly as normal status or 98.8 % of the data. While 47(1.16%) is miss classified. In other side 48 records as medium or 26.3 % correctly classified as medium.Whereas 134 or 73.6 % were in correctly or miss classified.

5.5.2. EXPERMENT 2

In this scenario by applying the two methods on the PART tree rule induction which is held for the under experiment one means (70/30) and 10 fold cross validation methods. The result of the experiment is like Table 5.9.

Table 5.9: Experiment result of PART rule induction algorithm with pruned.

Model	PART tree algorithm pruned	
	70/30 methods	By 10 fold cross validation
Instances in number	4218	4218
No of rules	317	317
Time (sec)	0.41 sec	0.62
Average Precision	0.958	0.956
Average ROC area	0.834	0.822
Average TP rate	0.954	0.958
Average TN rate	0.954	0.958
Average FP rate	0.427	0.511
Accuracy	95.415	95.753
Correctly Classified	1207	4039
Incorrectly Classified	58	179
Recall	0.954	0.95

From the above Table 5.9 the result of the experiment is conducted according to the selected methods which is (70/30) for training and testing.by the (70/30) the classifiers scored the accuracy of 95.415% which means from the data set of 4218 the classifiers classified 1207 instances correctly and 58 incorrectly. For the first method time taken to classify is almost **0.41** sec and the Roc area of 0.834.

For the second method means by 10 fold cross validation techniques the classifiers scored 95.753 % for the data set of 4218 instances. From this number of instances 4039 records are classified correctly and 179 records are incorrectly classified. Again on here from the two methods the 10 fold cross validation has a promising result than the (70/30).The confusion matrix looks like below for the 10 fold cross validation methods for pruned PART rule induction.

Table 5.10: The confusion matrix of PART tree rule induction for pruned

=== Confusion Matrix ===		
A	B	<-- classified as
3954	82	a=normal
97	85	b=medium

For the second experiment of this scenario the good result is produced by the second method means 10 fold cross validation. The result scored by using the two methods is 95.415 and 95.753 respectively. The above Table 5.10 is a confusion matrix of the 95.753 which is scored by pruned part tree by 10 fold cross validation. From the above Table 5.10 the among the record wich found on the dataset, 3954 records are correctly classified as normal and 85 records are correctly classified as medium. And 97 records are correctly classified as normal and 82 records correctly classified as medium. The confusion matrix is expanded according to the table below

Table 5.11: General confusing matrix of the pruned PART

	Predicting maternal mortality rate		
	Normal	Medium	Total
Maternal status as normal	3954	82	4036
Maternal status as medium	97	85	182
Total	4051	167	4218

For the above PART Table 5.11 again evaluating the model based on sensitivity and specificity are very significance for decision making. For that reason the result of the above confusion matrix indicate that the sensitivity of this test was (3954/4036) = 97.96 % and the specificity was (85/182) = 47.7%. The test indicates that the models appear to be decreasing more than the un pruned part algorithm. Because, based on the evaluation criteria, the classifier correctly classifies clients as Normal status rate with 97.96% accuracy but this is decreasing depend on previous result.

To summarize the above Table 5.11 from the whole data 3954 is classified as normal status or 97.96 % as a normal status while 82 records (2.03%) miss classified. Also 85 of the records or

46.70 % is classified as medium status. While 97 of the records or 53.29 % of the records is miss classified.

5.6. EXPERMENT WITH NAIVE BAYES MODEL BUILDING

Naïve Bayesian classifiers assume that there are no dependencies among the attributes. It has been called 'Naïve' due to the fact that it assumes mutually independent attributes. It uses a very intuitive technique which uses greatly simplify the design process. It does not require large amounts of data before learning can begin. This one of the reason that the researcher proved the 4218 instances are enough for the prediction of maternal mortality rate (Guven, 2012).

5.6.1. EXIPERMENT 1

The experiment of the naïve Bayes model building is performed now by the 70/30 training and testing data set. The 10 fold cross validation wich is done with the previous experiments of J48 and PART tree algorithm. Therefore the result of the naïve Bayes experiment is like the Table below.

Table 5.12: Experimental result of Naïve Bayes algorithm

Model	Naïve Bayes	
	70/30 methods	By 10 fold cross validation
Instances in number	4218	4218
Time (sec)	0.01sec	0.01sec
Average Precision	0.925	0.931
Average ROC area	0.755	0.768
Average TP rate	0.95	0.954
Average TN rate	0.905	0.904
Average FP rate	0.905	0.904
Accuracy	95.0198	95.377
Correctly Classified	1202	4023
Incorrectly Classified	63	195
Recall	0.05	0.954

The above Table 5.12 shows the result of the Naïve Bayes algorithm experiment which held on the 4218 dataset. This also experimented by two methods which is 70/30 and 10 fold cross validation. On the first methodology the classifiers scored the accuracy of 95.01 % and correctly classified the record 1202.additionally on this experiment 63 records are incorrectly classified. The second method which is 10 fold cross validation the result of the classifiers is more promised than the first method which held by (70/30). The classifiers scored the accuracy of 95.377 which is more accurate than the first method. Simultaneously the performance that scored both methods looks like relative but the ROC area, recall, precision and TP rate still the 10 fold cross validation is better. The time taken by both methods is still the same 0.01 sec.

In the above experiment the naïve Bayes algorithm scored a better accuracy by the 10 fold cross validation method. The confusion matrix of the naïve Bayes is like Table 5.13.

Table 5.13: The confusion matrix of the Naïve Bayes classifiers

=== Confusion Matrix ===		
A	B	<-- classified as
4013	31	a=normal
164	10	b=medium

The above confusion matrix which is in the Table of 5.13 is for the 10 fold cross validation which shows the better accuracy in the experiment. The classified and non-classified records are saved in the form of matrix. This shows from the 4218 data set 4013 records are correctly classified as normal and 10 records classified as medium, as well as 164 as normal and 31 as medium is correctly classified in the records. As a whole this shows how well the classifiers in the classification of instances as normal and medium for the prediction of maternal mortality rate.

Table 5.14: The general confusion matrix of Naïve Bayes algorithm

	Predicting maternal mortality rate		
	Normal	Medium	Total
Maternal status as normal	4013	31	4044
Maternal status as medium	164	10	174
Total	4177	41	4218

The previous PART rule induction Table 5.14 again evaluated based on sensitivity and specificity since it is very crucial for decision making. For that reason the result of the above confusion matrix indicate that the sensitivity of this test was (4013/4044) = 99.23 % and the specificity was (10/174) = 5.74 %. The test indicates that the models appear to be increased in sensitivity more than the un pruned PART algorithm. Because, based on the evaluation criteria, the classifier correctly classifies clients as Normal status rate with 99.23% accuracy but this is decreasing depend on previous result.

Lastly 4013 records or 99.23% records are correctly classified as normal status and 31 or 0.76 % is miss-classified. Also 10 (5.74%) records classified as medium status and 164 (94.25%) is miss classified.

5.7. COMPARISON OF J48 DECISION TREE, NAIVE BAYES AND PART RULE INDUCTION ALGORITHM

As observed in the previous sections, different experiments was held by (70/30) (training and testing) and 10 fold cross validation on the three selected model which could predict the maternal mortality i.e J48, PART and Naïve Bayes. But from the three algorithms Naïve Bayes scored the least accuracy regarding with performance and instance classification. The main aim of this study is to find a best performing model for the prediction of maternal mortality. Those three models identify the best attributes which answers the question of this study under chapter one of this study. The hybrid model is the main method of the researcher, due to this comparing and contrast the models performance is necessarily tested before developing the model for the stated problem.

In the previous section of this scenario the three selected model for this study scored different results with methodology of (70/30) and 10 fold cross validation. On this scenario the Naïve Bayes rule induction is the one which shows the lowest performing model, therefore comparing and contrasting the three model together is important for the completion of this study.

The important thing that was considered during comparison of models is the simplicity of the model for users and applicability of the model in actual working areas. Therefore, for the comparison of the models primarily the accuracy, precision, WTPR, WFPR, WROC, Recall of the classifier and the time required for building the models are taken. The details of comparison of classification models were summarized in Table 5.15.

Table 5.15: Comparison of the models.

Models	No of attributes	Time in sec	Instance classified	Accuracy %	WTPR	WFPR	WROC	precision	Recall
J48 pruned	14	0.14 Sec	4034	95.637	0.956	0.951	0.538	0.958	0.956
J48 un pruned	14	0.02sec	4165	98.7435	0.987	0.226	0.988	0.987	0.935
PART pruned	14	0.62	4039	95.753	0.958	0.511	0.822	0.956	0.95
PART un pruned	14	0.12 sec	4037	95.7	0.957	0.705	0.786	0.948	0.957
Naïve Bayes	14	0.01 sec	4023	95.377	0.95	0.904	0.768	0.931	0.95

The above Table 5.15 shows the comparison of the three models which has a good performing behavior based the result of the previous scenario. Among the result that the Table 5.15 holds the result of J48 decision tree algorithm with un pruned is the good performing one from the previous scenario. This is clearly shown in the Table 5.15 which is the classifiers accuracy is

98.74 % or 4165 were classified correctly from the dataset of 4218 records. The left 1.2% or 53 records are incorrectly classified. Even the time elapsed to classify those records is 0.02 sec which is very minimum time next to Naïve Bayes. Beside this the area of the ROC (Receiver Operating Characteristics) which proves how well the classifiers done is approximately very good. The ROC of these classifiers is 0.988 which is approximately one. The more the ROC area result is close to one or the more the area under the curve is wide the performance of the classifiers is good (Provost and Fawcett, 1998). Maimon, (2008) discusses how well the classifiers explain the performance of the model through TP and TN. On the diagonal the number of observations that have been correctly classified for each class; the off-diagonal elements indicate the number of observations that have been incorrectly classified or helps to see if the system is confusing which is discussed under previous experiments. Additionally under the scenario of J48 decision tree the number of leaves 1848 and the tree size is 1917 the result of this experiment. Those leaves and tree size is help full to produce the best rule which works for the maternal mortality rate prediction.This experiment also prone from error which means that the accuracy that is registered by this model is very high. Accuracy is a measure of the closeness to true value means 98.74% is good result based on the input inserted into J48 algorithm (Analysis of Errors, 2013).

The second algorithm which is next J48 algorithm for the prediction of maternal mortality rate is the PART rule induction algorithm as depicted by the Table 5.10.The algorithm performs classification performance of 95.7563% or 4039 records classified correctly and 4.2437% or 179 records are classified in correctly.As well as 317 rules also discovered for the purpose of predicting maternal mortality rate. To produce the above mentioned result it takes 0.62 sec of time elapse which is the greater time taken than the first experiment of J48 decision tree algorithm. In addition to these performance parameters, the model has generated a total of 317 rules to represent the patterns found within the dataset. The TP, ROC and TN are the same evaluation techniques which depends on the area means the one which has greater ROC area the one with performed well.

To sum up the above Table 5.15 which contains the result of three Algorithms, The J48 decision tree algorithm is the most performed model and PART tree rule induction is second and Naive Bayes is the third. The J48 decision tree result is the most performing model more than the rest

and the results initiates the accuracy and performance of model which predict the maternal mortality rate. Additionally the three models specificity and sensitivity showed according to the Table 5.16 below.

Table 5.16: Sensitivity and specificity result

Model	sensitivity	specificity
J48 pruned	97.4	43.2%
J48 un pruned	99.76%	76.4%
PART pruned	97.96%	47.7%
PART un pruned	98.86%	26.37%
Naïve Bayes	99.23%	5.74%

Again by summarizing the above Table 5.16 the sensitivity and specificity of the J48 un pruned has a maximum percent more than other which shows how much the model is approaching to classify the classification. This also proves that the J48 un pruned tree has more accurate than the rest of PART and Naïve Bayes (99.76%, 98.86%, 99.23%) respectively. On here the Naïve Bayes has a good sensitivity but very minimum specificity, due to this the classifiers is not accurate.

5.8. GENERATING RULES FROM DECISION TREE

After consecutive experiments in building the best decision tree model, the next step is to generate, rules by tracing through the branches up to leafs. In the previous experiment the J48 algorithm has a pretty good result more than PART and Naïve Bayes algorithm. A rule is a correlation found between the main variable (dependent) and the others (independent).The corresponding rules extracted form decision trees is listed below and Some of the rules believe to be interesting are randomly selected and presented as follows. Additionally, the rule is selected by discussing with domain experts weather the rule discovered is valid and help full for the prediction of maternal mortality rate.

Rule 1:

If Mothers BP=Normal BP and Diagnosis=Before Delivery and Apgar score=6&8 and Mothers Age=18-20 And Length of Stay =1 and Address =Inner city or Debub or Rural area or other THEN Normal (9.0).

Rule 2:

If Mothers BP=Normal BP and Diagnosis=Before Delivery and APGAR Score= 7&9 or 8&8 or 7&8 and Mothers Age=18-20 And Length of Stay =1 and Address =Inner city or Debub or Rural area or other and Indication=Normal Labor or Poor Maternal Effort THEN Normal (2.0).

Rule 3:

If Mothers BP=Normal BP and Diagnosis=Before Delivery and APGAR Score= 6&8 and Mothers Age=18-20 And Length of Stay =1 and Address =Inner city or Debub or Rural area or other and Indication=Brady Cardia THEN Medium (1.0).

Rule 4:

If Mothers BP=High BP and Diagnosis=After Delivery and APGAR Score= 0 and Mothers Age=27-43 And Length of Stay =1-4 and Address =Inner city or Debub or Rural area or other and Indication=Normal Labor or Poor Maternal Effort or Normal Labor THEN Medium (1.0).

Rule 5:

If Mothers BP=Normal BP and Diagnosis=Before Delivery and APGAR Score= 0 and Mothers Age=17-23 And Length of Stay =1-34 and Address = Rural area and Indication=Poor Maternal Effort THEN Medium (5.0).

Rule 6:

If Mothers BP=Normal BP and Diagnosis=Before Delivery and APGAR Score= 8&9 and Mothers Age=23-43 And Length of Stay =1-4 and Address =Inner city and Indication=Normal Labor THEN Normal (1.1.0).

Rule 7:

If Mothers BP=High BP and Diagnosis=After Delivery and APGAR Score=9&6 and Mothers Age=30-43 And Length of Stay =6-34 and Address =Debub or Rural Area and Indication=Poor Maternal Effort THEN Medium (1.2.0).

Rule 8:

If Mothers BP=Normal BP and Diagnosis=Before Delivery and APGAR Score= 6&9 or 7&8 or 8&8 or 8&9 or 8&7 and Mothers Age=23-43 And Length of Stay =1-34 and Address =Rural Area and Indication=Brady Cardia THEN Medium (3.0).

Rule 9:

If Mothers BP=Normal BP and Diagnosis=After Delivery and APGAR Score= 6&8 and Mothers Age=17-20 And Length of Stay =1-34 and Address =Inner City and Indication=Normal Labor THEN Normal (1.0).

Rule 10:

If Mothers BP=Normal BP and Diagnosis=After Delivery and APGAR Score= 6&8 or 7&9 or 7&8 or 8&8 And Length of Stay =1-34 and Address =Inner City or Rural Area or Debub or Other and Indication=Poor Maternal THEN Normal (6.0).

Rule 11:

If Mothers BP=High BP and Diagnosis=After Delivery and APGAR Score= 0 And Length of Stay =20-34 and Mothers Age=25-43 and Address = Rural Area and Indication=Poor Maternal THEN Medium (3.0).

Rule 12:

APGAR Score= 6&8 or 7&9 or 7&8 or 8&8 and Mothers Age=17-23 and Mothers BP=High BP and Indication=Normal Labor and Diagnosis=Before Delivery and Length of Stay =2 and Address=Inner City or Rural Area or Other THEN Normal (27.04)

Rule 13:

APGAR Score= 6&8 or 7&9 or 7&8 or 8&8 and Mothers Age=17-43 and Mothers BP=Medium BP and Indication=Normal Labor and Diagnosis=Before Delivery and Length of Stay =2 and Address=Inner City or Rural Area or Other or Debub THEN Normal (8.02)

Rule 14:

Diagnosis=After Delivery and Length of Stay =<1 and Mothers Age=17-27 Address=Inner City or Rural Area or Other or Debub THEN Normal (3.0/1.0)

Rule 15:

APGAR Score=6&6 and Length of Stay=1-4 and Adress=Rural Area or Other or Debub or Inner City THEN Normal (0.0)

Rule 16:

Mothers BP= Medium BP and Indication = Poor maternal and Condition on Discharge = Good THEN Normal (51.0)

Rule 17:

Mother BP=High BP and Indication = Brady Cardia THEN Medium (1.0)

Rule 18:

Indication = Poor Maternal Effort and Condition on Discharge =Good and Mothers BP=Medium BP THEN Normal (1.0)

Rule 19:

Indication=Brady Cardia and Condition on Discharge =Not Good and APGAR Score=6&8 or 7&9 or 8&8 and Mothers BP = medium THEN Medium (0.0)

5.9. RESULT AND DISCUSSION

The purpose of this research is to predict maternal mortality rate using data mining techniques and finding attributes which have strong relationship with woman's during pregnancy period and to develop predicting the future life status of initiation time predictive model, the findings are discussed in this section. The main aim of data mining is classifying the attribute based on the given attribute. This is achieved by decision trees even though three algorithms are selected for this purpose. The algorithm tries to discover relationships between the attributes that would make it possible to predict the outcome (Fabricio and Leonardo, 2001). According to the world health statics, (2014) the distance and lack of basic infrastructure is the most affecting problem of pregnant women's future life.This include mothers BP and the distance of health station which the pregnant woman comes from is under this category. From the maternity ward data the J48 algorithm is discovering this attributes which validates the factors for the death of pregnant woman if they are not counseled.

The rule that is discovered from Decision tree section is the attempt of finding the attribute which predict the maternal mortality rate. For this sake the previous rule which mined out by the un pruned J48 model is the result of this study which correct out the status and future conditions of mothers which comes to the Maternity ward for the service of labor and delivery.

On this study different types of scenario were conducted for the purpose of maternal mortality prediction, as well which attribute, which model and which algorithm would perform very well and it is approved under this study. For this study three algorithm were selected to test on the maternal mortality rate i.e PART, J48 and Naïve Bayes algorithms. Three of them are raised under the methodology of this study. Therefore, analyzing one by one and seeing the result that they performed during the previous experiment has been tabularized accordingly. Additionally the J48 algorithm is the most performing model more than the rest of the algorithm. And the other algorithm had been resulted according to the nature and ability of evaluation based on the natural content of the parameter means how the algorithm is set by default.

The J48 algorithm is the most accurate model from the other due the result that this algorithm displayed in case of performance, time, sensitivity, specificity and Matrix. From the previous scenario the J48 algorithm had scored a time of 0.02 seconds to classy the 4165 records according the class they belongs too. Beside this, this model also showed the good performance

more than the other. The ROC which this model displayed is almost approximate to one which is 0.988 and the result of precision and recall (0.987 and 0.935) also pretty good more than the left model. This model showed the most performing one in case of predicting the maternal mortality rate status. This is proved on the Table 5.4 result of J48 model. This model scored the accuracy of 98.74% to classify the data. The result is without any bias means the specificity and sensitivity that this model scored proves how well the J48 algorithm is predicting maternal mortality rate. The specificity and sensitivity that they score is 99.96% and 76.4% respectively. This model proves by classifying the attributes which predicts the futurity of maternal status by classifying the result as Normal and Medium. The rule that is generated from the J48 algorithm is the best rule wich solve the problem of maternal mortality problem since the rule supports the attribute which determine women's factors accordingly. Decision Trees are built from nodes, branches and leaves that indicate the variables, conditions, and outcomes, respectively. The most predictive variable is placed at the top node of the tree based on the ID3 or C4.5 algorithms (Quinlan, 1993). Mother BP is the most predictive variables from the data set .

The second most performing model is the PART pruned model which is the second one according to the above criteria. This model performed the second promising result next to the J48 algorithm. This model scored the 95.7% accuracy on the general data to classify the status of maternal mortality rate prediction. The time taken to perform the general data by this algorithm is 0.12 sec and to classify the 4034 instances of the records. The recall (0.957) and precision (0.948) with the ROC area of 0.786. This result is the most promising result next to J48 algorithm by understanding the experiment result of the model. The specificity and sensitivity of this model is 98.86% and 26.37 %respectively. The more the sensitivity is decreasing the more the accuracy of the model is decreasing. Also, the specificity and sensitivity also depends on the number of correctly classified i.e in this case Normal and Medium class which is very important class for the prediction of maternal mortality rate. To sum up about this model, the model is very promising model next to J48 algorithm depending on the Table 5.6 of the previous experiment.

The third model which selected for this study is the Naïve Bayesian algorithm which is almost very close to the PART un pruned. This model scores the accuracy of 95.37% by classifying 4023 records. The time taken to display this result is 0.01sec which is good time but poor

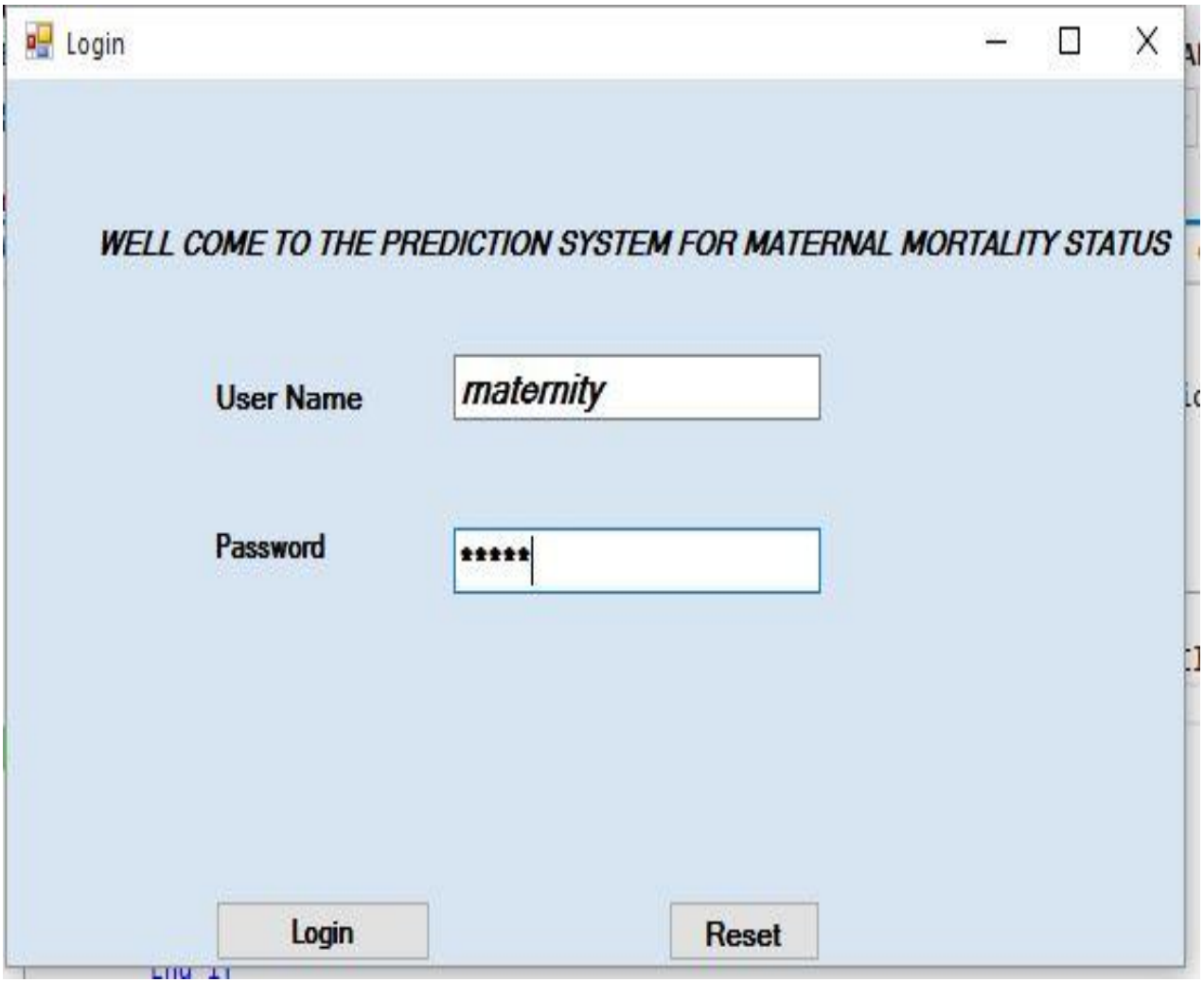
performance by comparing the other model. The sensitivity and specify of this model is 99.23% and 5.74% respectively. This is the third poor result depending on the previous result.

To generalize, the J48 model is the most performing model with a good accuracy of results. The PART rule induction is the second most performing model next to J48 model whereas the Naïve Bayes is the last poor one. Among this the J48 algorithm is the good performing model which shows the status of maternal mortality rate by the classifying as normal and Medium which is proved by the confusion matrix of the J48 algorithm.

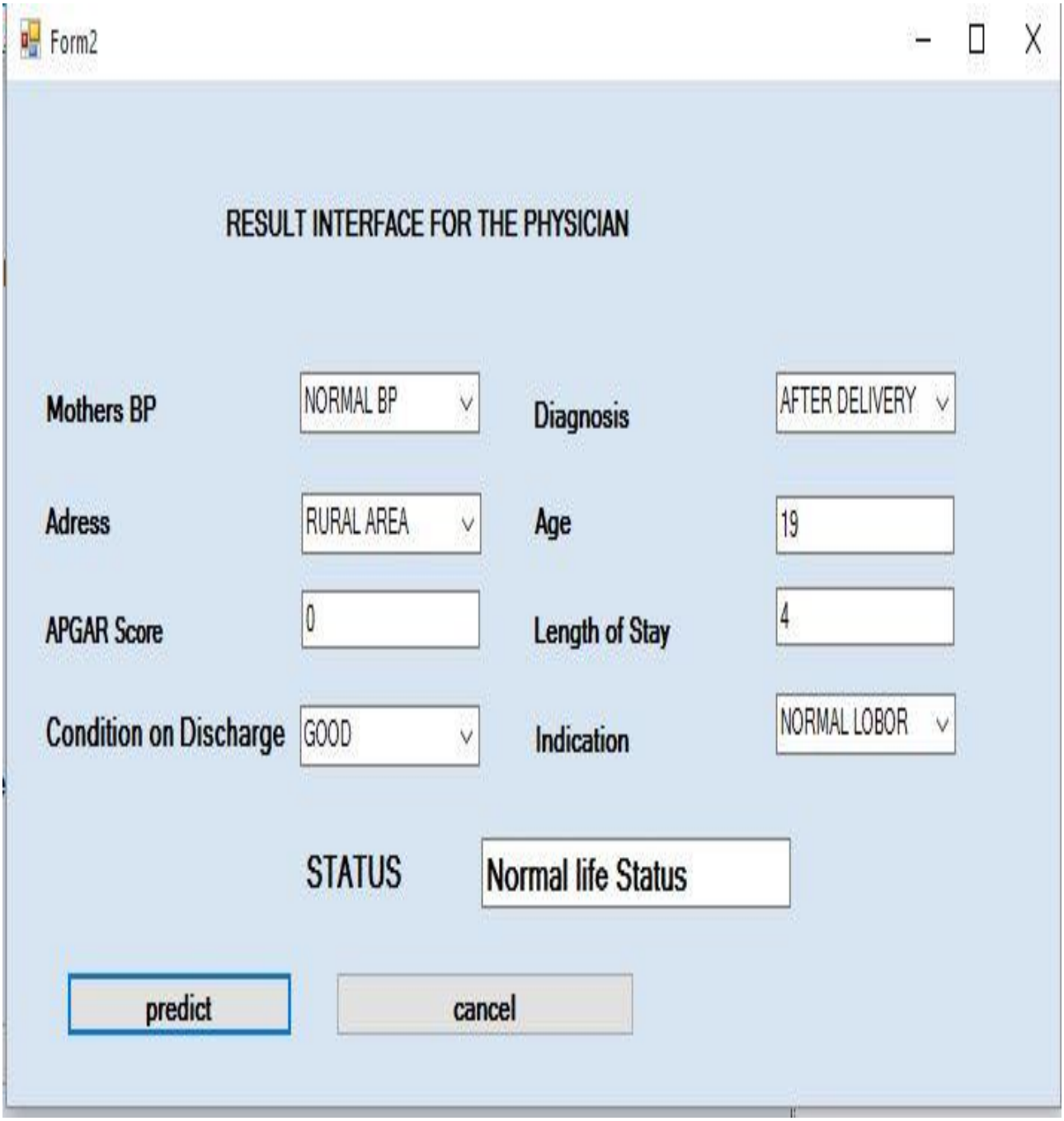
The result which found on this study is the J48 algorithm has the most performing model more than the PART and Naïve Bayes. Beside this the J48 decision tree discovered the 19 rules which and identifying the attribute which might predict the maternal mortality rate status by discovering the most important knowledge with the most important attributes which could determine maternal mortality i.e "Mothers BP", "Address", "APGAR Score"," Diagnosis", "Age", "Length of stay" and "Indication". Taking the first Rule obtained from classification algorithm of J48, the future life status of mothers after they had service from the maternity ward is Normal future status of mother after the service. This indicates that a pregnant mother beginning with such socio-demographic, clinical and biological parameters has a confidence of Normal life conditions with or without her baby. If this condition not fulfilled the result is inverse which means according the other rule like Rule No 8 are which makes the life of mothers under risky after she had a service from the maternity Wards. Shortly the physicians would depend on the model and give counseling for the pregnant mother after delivery. This means that if any woman comes and had delivery, then affected by Rule No 8, she needs special attention from the physician to save her life from risk conditions which could be loss of her life or death. Abdella, (2010) explains about the factors which could determine pregnant women's life from delay one up to delay three which considers nutrition with mothers blood pressure and communication network and address of pregnant woman. This is to strengthen the rule and attribute which is discovered by J48 decision tree could determine the status of pregnant women's life after she serviced from the healthcare.

5.10. DEVELOPING A PROTOTYPE

The final objective of this study was developing a prototype interface that assists physician easy access to the identified knowledgebase. The final selected if-then rules are used to implement the selected best models. The programming tool used to host the identified rules is Microsoft visual basic 2012. Therefore, only those rules which are suggested to be important by domain experts are placed into this prototype which means all the rules for predicting Maternal Mortality rate and any question which is out of this rule is not answered by this system. The following figure is the main graphical user interface used to run the commands to predict maternal mortality rate



5.4: The login interface for the prepared system



5.5: Result interface for the prediction

CHAPTER SIX

CONCLUSION AND RECOMMENDATION

6.1. CONCLUSION

Data mining techniques can be applied to maternity ward database to predict the maternal mortality rate and helps to discover links or patterns in a store of data. It can help to discover interesting associations between data items of pregnant mother's records and enable to predict missing or unknown values based on rules mined through the process of Decision tree rules with classification mining.

In order to achieve the objective of the study the researcher used three popular classification algorithms (J48, Naïve Bayes and PART rule induction) and long process of data cleansing, data and dimensionality reduction and transformation used it to build the prediction models on 4218 instances and 14 attribute of maternity ward dataset from Jimma university specialized Hospital.

The result from the three applied classification algorithm of the experiment the J48 un pruned decision tree had scored the pretty good results for the prediction of maternal mortality rate by classifying the attribute which could predict mothers future life status. This is already proved by classifying the last result as normal and Medium which means as Normal conditions pregnant mothers life could not fail under risk whereas under medium class the life of woman's will fail under risk, so additional counseling and other important advice should be given by the physicians.

The study has shown the necessity to experiment as many classification algorithms as possible before selecting and using a single algorithm for prediction. Eight attribute is selected from fourteen to predict the maternal mortality rate. The un pruned J48 algorithm is a good model for the prediction of maternal mortality rate.

6.2. RECOMMENDATION

In this research work, efforts have been made to apply data mining technology to predict maternal mortality rate using J48, Naïve Bayes and PART rule induction tree algorithms.

Thus, based on the result of these research future works are lined. This would more enhance applicability of data mining technology in maternal mortality status so the death of maternal is determined by the developed prototype and attributes. This is in lined in health prevention and control activities with advocacy efforts of maternal mortality reduction policy in rural communities of the country. The following recommendation is lined:

- Beside this instead of developing the model by using and testing three algorithms adding the algorithms more than three would provide a more accurate result like Neural network, since neural network is used in many health related research.
- In both the decision tree and Naive Bayes approaches the result is an encouraging output, still performance improvement is expected.
- Other classification algorithms such as Neural networks and Bayesian network (Belief network) which have also been proved to be important techniques in the health care sector could be applied by using the entire dataset.
- It is appropriate to predict the survival years of the individual in the area corresponding to sample data available through data mining technology and it is also possible to guess the life expectancy of the pregnant woman would live after she served from maternity ward.
- The decision tree and Naïve Bayes reported promising results and hence they could be applied in the area of maternal mortality rate predictive modeling, decision tree tends to perform better.
- Thus, it would be more optimal for the Jimma University Specialized hospital maternity ward to employ the developed model with this technique.

REFERENCES

Abdella, A. (2010). Maternal Mortality Trend In Ethiopia. *Health Dev,Special Issue 1*, 116-119.

Abraham, T. (2005). Application of Data Mining Techniques To Discover Cause of Under-Five Children Admission To Pediatric Ward:The Case of Nigist Eleni Mohammed Memorial Zonal Hospital.Masters Thesis.Addis Abeba Unversity:Addis Ababa,Ethiopia.

Admasu, K. (2014). Supporting Evidence-Based Decision Making:To Wards The Achievement of The Health Millennium Development Goals. *Policy And Practice Information For Action.*, 3-4.

Agarwal, D. T. (2013). A Survey On Data Mining Approaches For Healthcare. *International Journal of Bio-Science And Bio-Technology Vol.5, No.5* , 241-266.

Ana , D. W. (2005). National Estimates For Maternal Mortality: An Analysis Based On The Who Systematic Review of Maternal Mortality And Morbidity. *Bmc Public Health*, 2-6.

Shagaw,A. (2002). Application of Data Mining Technology To Predictchild Mortality Patterns.The Case of Butajira Rural Health Project:Masters Thesis,Addis Ababa Unvesrity,Addis Abeba,Ethiopia.

Selam, A. (2011). Predicting The Occurrence of Measles Outbreak In Ethiopia Using Data Mining Technology:Un Published Masters Thesis.Addis Abeba.Addis Abeba Unversity.

Ayale, D. (2013). Predicting Maternal Health Care Seeking Pattern.Masters Thesis.Addis Ababa Univesity,Addis Ababa,Ethiopia. 69-71.

Chaudhary, P. (2015). Data Mining System, Functionalities And Applications,A Radical Review. *International Journal of Innovations In Engineering And Technology (Ijiet),Vol 5*, 449.

Danso, S. O. (2006). An Exploration of Classification Prediction Techniques In Data Mining:The Insurance Domain. 6-29.

Delen, D. O. (2008). *Advanced Data Mining Techniques.* Heidelberg: Springer.

Dharani Shree, K. (2014). Application of Data Mining Techniques In Health Care Industry. *International Journal Of Computer Science And Information Technology Research*, 23-25.

Getachew, D. (2013).Application of Data Mining Techniques To Predict Antiretroviral Therapy Initation Time The Case of Adama And Ambo Hospitals, Oromia Regional State.Masters Thesis.:Addis Abeba Unversity:Addis Abeba.Ethiopia.

Gheware, A. K. (2014). Data Mining Task, Tools, Techniques And Applications. *International Journalo of Advanced Research In Computer And Communication Engineering*, 8095.

Goel, M. G. (2015). Data Mining - Techniques, Methods And Algorithms: A Review on Tools And Their Validity. *International Journal of Computer Applications,Vol 113 – No. 18*, 1-8.

Hailemariam, T. (2012). Application of Data Mining For Predicting Adult Mortality.Masters Thesis.Addis Abeba,Addis Abeba Unversity:Ethiopia.

Han, J. (2006). *Data Mining Concepts And Techniques.* Singapore,Sydne: Morgan Kaufmann Publisher.

Ian.H.Witten, E. F. (2011). *Data Mining : Practical Machine Learning Tools And Techniques.* Newyork: Morgan Kaufmann .

Jackson, J. (2002). Data Mining: A Conceptual Overview. *Communications of The Association For Information Systems (Vol 8*, 267-296).

Jyoti Soni, U. A. (2011). Predictive Data Mining For Medical Diagnosis: An Overview of Heart Disease Prediction. *International Journal Of Computer Applications (0975 – 8887) Vol 17– No.8,*, 5-47.

Kale, T. D. (2015, Dece). *Application of Data Mining Techniques to Discover Cause Of Underfive Children Admission to Pediatric Ward: The Case of Nigist Eleni*. Retrieved From Http://Dx.Doi.Org/10.4172/2157-7420.1000178.

Kale, T. (2015). Application of Data Mining Techniques to Discover Cause of Underfive Children Admission To Pediatric Ward: The Case of Nigist Eleni . *J Health Med Inform*, 10-14.

Kamber, J. H. (2006). *Data Mining Concept And Techniques.* Amsterdam Boston: Morgan Kaufmann.

Kenneth Hill, C. A. (2001). Estimates of Maternal Mortality For 1995. *Bulletin of The World Health Organization*, 79.

Krzysztof J. Cios, W. P. (2007). *A Knowledge Discovery Approach.* New York: Springer Science+Business Media, Llc, 233 Spring Street.

Kumar R, K. A. (2012). A Modified Tree Classification In Data Mining,Vol 12. *Global Journals Inc*, 58-63.

Kumar, Y. A. (2012). Analysis of Bayes, Neural Network and Tree Classifier of Classification Technique In Data Mining Using Weka:. *Cs & It05*, 359– 369.

Maimon, L. R. (2008). *Data Mining With Decision Trees:Theories And Application,Vol 69.* Singapore: World Scientific Publishing Co. Pte. Ltd.

Margaret , K. J. (2010). Maternal Mortality For 181 Countries, 1980–2008: . P. 1.

Maurizio, M. (2011). Data Mining Concepts And Techniques. *E-Commerce Winter*, 23-34.

Minale Tefera, M. M. (2014). Application of Data Mining Techniques to Predict Urinary Fistula Surgical Repair Outcome: The Case of Addis Ababa Fistula Hospital, Addis Ababa, Ethiopia. *Health & Medical Informatics*, 23-37.

Mulugeta, T. (2013). Constructing A Predictive Model For Occurrence of Tuberculosis: The Case of Menelik Ii Hospital And St. Peters TB Specialized Hospital.Masters Thesis.Adis Abeba Unversity. Addis Abeba, Ethiopia.

Obenshain, M. (2004). Application of Data Mining Techniques to Healthcare Data. *Infection Control And Hospital Epidemiology, Vol. 25*, 690-695.

Parvez, A. (2015).Techniques of Data Mining In Healthcare: A Review. *International Journal of Computer Applications ,Vol 120 ,No.15*, (0975 – 8887).

Paulraj. (2015). Prediction of Low Birth Weight Infants And Its Risk Factors Using Data. *Proceedings of The 2015 International Conference On Industrial Engineering And Operations Managemen*, 2.

Petri, C. (2010). Decision Trees. *Journal of Computer Enginering*, 23-56.

Pradhan, M. (2014). Data Mining& Health Care: Techniques of Application. *International Journal of Innovative Research In Computer And Communication Engineering Vol. 2, Issue 12, December* , 235-435.

Ranjani, M. D. (2013). Data Mining Applications In Healthcare Sector: A Study. *International Journal Of Scientific & Technology Research Vol 2, Issue 10, October 2013* , -.

Sivanandam, S. S. (2006). *Introduction To Data And Mining And Its Applications.* Berlin: Springer.

Srideivanai, C. (2014). Data Mining Techniques For Performance Evaluation of Diagnosis In Gestational Diabetes. *International Journal Of Current Research and Reveiew,Vol 2,No 10*, 91-98.

Sundar, A. (2012). Performance Analysis of Classification Data Mining. *International Journal of Engineering Science & Advanced Technology,Vol 2,No 2*, 471-475.

Venkateswaran, A. (2012). An Approach of Data Mining For Predicting The Chances . *The International Conference on Communication, Computing And Information Technology (Icccmit) 2012*, 1.

Wang, L. G. (nd). Data Mining Techniques for Mortality at Advanced Age. 3-27.

Witten, I. H. (2005). *Data Mining - Practical Machine Learning Tools and Techniques With Java Implementations.* San Francisco. Morgan Kauffmann Publishers.

APPENDIX 1: J48 10 FOLD CROSS VALDIATION RESULT

```
=== Run information ===
Scheme:       weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:     Maternity ward database-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-
Rfirst-last-weka.filters.unsupervised.attribute.NumericToNominal-Rfirst-last-
weka.filters.unsupervised.attribute.RemoveUseless-M99.0-
weka.filters.unsupervised.instance.Resample-S1-Z100.0-
weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-
weka.filters.unsupervised.attribute.NumericToNominal-Rfirst-last-
weka.filters.unsupervised.instance.Resample-S1-Z100.0
Instances:    4218
Attributes:   14
Test mode:    10-fold cross-validation
=== Classifier model (full training set) ===
J48 pruned tree
------------------
: Normal (4218.0/185.0)
Number of Leaves:                 1
Size of the tree:                 1
Time taken to build model: 0.14 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances        4034               95.6377 %
Incorrectly Classified Instances      184                4.3623 %
Total Number of Instances             4218
=== Detailed Accuracy By Class ===
               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 1         0.995      0.956     1         0.978      0.538    Normal
                 0.005     0          1         0.005     0.011      0.538    Medium
Weighted Avg.    0.956     0.951      0.958     0.956     0.935      0.538


=== Confusion Matrix ===
   a    b   <-- classified as
 4033   0 |   a = Normal
  184   1 |   b = Medium
```

APPENDIX 2: J48 UN PRUNED TREE 10 FOLD CROSS VALIDATION

Time taken to build model: 0.02 seconds

=== Evaluation on training set ===

=== Summary ===

Correctly Classified Instances	4165	98.7435 %
Incorrectly Classified Instances	53	1.2565 %
Kappa statistic	0.8334	
Mean absolute error	0.0199	
Root mean squared error	0.0993	
Relative absolute error	24.0069 %	
Root relative squared error	48.892 %	
Coverage of cases (0.95 level)	99.6207 %	
Mean rel. region size (0.95 level)	52.5367 %	
Total Number of Instances	4218	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.998	0.236	0.989	0.998	0.993	0.988	Normal
	0.764	0.002	0.933	0.764	0.84	0.988	Medium
Weighted Avg.	0.987	0.226	0.987	0.987	0.987	0.988	

=== Confusion Matrix ===

a	b	<-- classified as
4026	10	a = Normal
43	139	b = Medium

APPENDIX 3: THE PART ALGORITHM RESULT

=== Run information ===

Scheme: weka.classifiers.rules.PART -M 2 -C 0.25 -Q 1
Relation: Maternity ward database-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-weka.filters.unsupervised.attribute.NumericToNominal-Rfirst-last-weka.filters.unsupervised.attribute.RemoveUseless-M99.0-weka.filters.unsupervised.instance.Resample-S1-Z100.0
Instances: 4218
Attributes: 14
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

PART decision list

: Normal (5.03/1.0)
Number of Rules : 105

Time taken to build model: 0.35 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	4037	95.7089 %
Incorrectly Classified Instances	181	4.2911 %
Kappa statistic	0.3266	
Mean absolute error	0.0617	
Root mean squared error	0.1919	
Relative absolute error	74.5017 %	
Root relative squared error	94.4427 %	
Coverage of cases (0.95 level)	97.89 %	
Mean rel. region size (0.95 level)	57.3969 %	
Total Number of Instances	4218	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.988	0.736	0.967	0.988	0.978	0.786	Normal
	0.264	0.012	0.505	0.264	0.347	0.786	Medium
Weighted Avg.	0.957	0.705	0.948	0.957	0.951	0.786	

=== Confusion Matrix ===

a	b	<-- classified as
3989	47	a = Normal
134	48	b = Medium

APPENDIX 4: PART WITH CROSS FOLD VALIDATION WITH 70/30 METHOD

=== Run information ===

Scheme: weka.classifiers.rules.PART -M 2 -C 0.25 -Q 1
Relation: Maternity ward database-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-weka.filters.unsupervised.attribute.NumericToNominal-Rfirst-last-weka.filters.unsupervised.attribute.RemoveUseless-M99.0-weka.filters.unsupervised.instance.Resample-S1-Z100.0
Instances: 4218
Attributes: 14
Test mode: split 70.0% train, remainder test

=== Classifier model (full training set) ===

PART decision list

Mothers BP = High BP: Normal (970.0/1.0)

Number of Rules : 105
Time taken to build model: 0.14 seconds

=== Evaluation on test split ===
=== Summary ===

Correctly Classified Instances	1210	95.6522 %
Incorrectly Classified Instances	55	4.3478 %
Kappa statistic	0.2092	
Mean absolute error	0.0701	
Root mean squared error	0.1933	
Relative absolute error	84.8188 %	
Root relative squared error	95.638 %	
Coverage of cases (0.95 level)	98.1028 %	
Mean rel. region size (0.95 level)	56.3636 %	
Total Number of Instances	1265	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.993	0.852	0.963	0.993	0.978	0.793	Normal
	0.148	0.007	0.471	0.148	0.225	0.793	Medium
Weighted Avg.	0.957	0.816	0.942	0.957	0.946	0.793	

=== Confusion Matrix ===

a	b	<-- classified as
1202	9	a = Normal
46	8	b = Medium

APPENDIX 5" PART WITH 10 FOLD CROSS VALIDATION METHOD

=== Run information ===

Scheme: weka.classifiers.rules.PART -M 2 -C 0.25 -Q 1

Relation: Maternity ward database-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-weka.filters.unsupervised.attribute.NumericToNominal-Rfirst-last-weka.filters.unsupervised.attribute.RemoveUseless-M99.0-weka.filters.unsupervised.instance.Resample-S1-Z100.0

Instances: 4218

Attributes: 14

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

PART decision list

: Normal (5.03/1.0)

Number of Rules : 105

Time taken to build model: 0.12 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	4037	95.7089 %
Incorrectly Classified Instances	181	4.2911 %
Kappa statistic	0.3266	
Mean absolute error	0.0617	
Root mean squared error	0.1919	
Relative absolute error	74.5017 %	
Root relative squared error	94.4427 %	
Coverage of cases (0.95 level)	97.89 %	
Mean rel. region size (0.95 level)	57.3969 %	
Total Number of Instances	4218	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.988	0.736	0.967	0.988	0.978	0.786	Normal
	0.264	0.012	0.505	0.264	0.347	0.786	Medium
Weighted Avg.	0.957	0.705	0.948	0.957	0.951	0.786	

=== Confusion Matrix ===

a	b	<-- classified as
3989	47	a = Normal
134	48	b = Medium

Appendix 6: Pruned PART by 10 Fold Cross Validation.

=== Run information ===

Scheme: weka.classifiers.rules.PART -U -M 2 -C 0.25 -Q 1
Relation: Maternity ward database-weka.filters.unsupervised.attribute.Discretize-B10-M-1.0-Rfirst-last-weka.filters.unsupervised.attribute.NumericToNominal-Rfirst-last-weka.filters.unsupervised.attribute.RemoveUseless-M99.0-weka.filters.unsupervised.instance.Resample-S1-Z100.0
Instances: 4218
Attributes: 14
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

PART decision list

Number of Rules : 317

Time taken to build model: 0.62 seconds
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	4039	95.7563 %
Incorrectly Classified Instances	179	4.2437 %
Kappa statistic	0.465	
Mean absolute error	0.0481	
Root mean squared error	0.1887	
Relative absolute error	58.1403 %	
Root relative squared error	92.8791 %	
Coverage of cases (0.95 level)	97.724 %	
Mean rel. region size (0.95 level)	52.9279 %	
Total Number of Instances	4218	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.98	0.533	0.976	0.98	0.978	0.822	Normal
	0.467	0.02	0.509	0.467	0.487	0.822	Medium
Weighted Avg.	0.958	0.511	0.956	0.958	0.957	0.822	

=== Confusion Matrix ===

a	b	<-- classified as
3954	82	a = Normal
97	85	b = Medium

APPENDIX 7: DECISION TREE RESULT

Mothers BP = Normal BP
| Diagnosis = Before Delivery
| | Length of Stay = 1.0
| | | APGAR SCORE = 6&8
| | | | Mothers Age = '(-inf-17.3]': Normal (0.0)
| | | | Mothers Age = '(17.3-20.6]'
| | | | | Baby weight = 2.8: Normal (0.0)
| | | | | Baby weight = 3.5: Medium (2.0)
| | | | | Baby weight = 2.6: Normal (0.0)
| | | | | Baby weight = 3.4: Normal (0.0)
| | | | Baby weight = 3.0: Normal (1.0
| | | | Baby weight = 1.2: Normal (0.0)
| | | | Baby weight = 2, 2: Normal (0.0)
| | | | Baby weight = 4.9: Normal (0.0)
| | | | Baby weight = 4.6: Normal (0.0)
| | | APGAR SCORE = 9&9: Normal (1.0)
| | | APGAR SCORE = 6&9: Normal (70.12)
| | | APGAR SCORE = 8&10
| | | | Adress = Inner City: Normal (2.01/1.0)
| | | | Adress = Rural Area: Medium (3.0/1.0)
| | | | Adress = Rural Area : Medium (0.0)
| | | | Adress = Debub: Medium (0.0)
| | | | Adress = Other: Medium (0.0)
| | | | Adress = Rural Area : Medium (0.0)
| | | | Adress = Inner City : Medium (0.0)
| | | APGAR SCORE = 5&9: Normal (1.0)
| | | APGAR SCORE = 3&4: Normal (3.01)
| | | APGAR SCORE = 4&6: Normal (0.0)
| | | APGAR SCORE = 6&7

| | | | Mothers Age = '(-inf-17.3]': Normal (0.0)

| | | | Mothers Age = '(17.3-20.6]': Normal (18.08)

| | | | Mothers Age = '(20.6-23.9]': Normal (11.0)

| | | | Mothers Age = '(23.9-27.2]'

| | | | | Baby weight = 2.8: Normal (0.0)

| | | | | Baby weight = 3.5: Normal (0.0)

| | | Baby weight = 5.5: Normal (0.0)

| | | Baby weight = 3.12: Normal (0.0)

| | | Baby weight = 4.33: Normal (0.0)

| | APGAR SCORE = 5&7: Normal (0.0)

| | APGAR SCORE = 7&6: Normal (29.07)

| | APGAR SCORE = 5& 7: Normal (0.0)

| | | APGAR SCORE = 8& 10: Normal (0.0)

| | | APGAR SCORE = 8&9`: Normal (0.0)

| | | APGAR SCORE = 7&5: Normal (0.0)

: Normal (0.0)

| | | APGAR SCORE = 7&7: Normal (0.0)

| | | APGAR SCORE = 6&10: Normal (0.0)

| | | APGAR SCORE = 7&9`: Normal (0.0)

| | | APGAR SCORE = 8&6: Normal (0.0)

| | | APGAR SCORE = 9&6: Normal (0.0)

| | | APGAR SCORE = 9&7: Normal (0.0)

| | | APGAR SCORE = 6&4: Normal (0.0)

| | | APGAR SCORE = 10&7: Normal (1.0)

| | | APGAR SCORE = 10&8: Normal (0.0)

| | | APGAR SCORE = 10&6: Normal (0.0)

Mothers BP = Medium Bp: Normal (1.0)

Number of Leaves : 1848

Size of the tree : 1917

APPENDIX 8: CODE FOR THE LOGIN PAGE

```
Private Sub Button1_Click(sender As Object, e As EventArgs) Handles Button1.Click

        If TextBox1.Text = ("maternity") And TextBox2.Text = ("12345") Then
            Form2.Show()
            Me.Hide()
        Else
            MsgBox("Incorrect Password or User Name")
        End If

    End Sub

    Private Sub TextBox1_TextChanged(sender As Object, e As EventArgs) Handles
TextBox1.TextChanged
        Dim Username, Password, Login As New Double
        Username = Val(TextBox1.Text)
        Password = Val(TextBox2.Text)

    End Sub

    Private Sub Button2_Click(sender As Object, e As EventArgs) Handles Button2.Click
        TextBox1.Clear()
        TextBox2.Clear()
    End Sub
End Class
```

APPENDIX 9: CODE FOR THE FORM TWO

```
Private Sub Button1_Click(sender As Object, e As EventArgs) Handles Button1.Click


        If ComboBox1.Text = "NORMAL BP" And ComboBox5.Text = "BRADY CARDIA" And
ComboBox4.Text = "BEFORE DELIVERY" And ComboBox2.Text = "RURAL AREA" And TextBox2.Text =
"0" And ComboBox3.Text = "GOOD" And TextBox3.Text = "19" And TextBox4.Text = "4" Then
            TextBox1.Text = "Risky life status"
        ElseIf ComboBox1.Text = "NORMAL BP" And ComboBox5.Text = "NORMAL LOBOR" And
ComboBox4.Text = "BEFORE DELIVERY" And ComboBox2.Text = "RURAL AREA" And TextBox2.Text =
"0" And ComboBox3.Text = "GOOD" And TextBox3.Text = "19" And TextBox4.Text = "4" Then
            TextBox1.Text = "Risky life status"
        ElseIf ComboBox1.Text = "NORMAL BP" And ComboBox5.Text = "NORMAL LOBOR" And
ComboBox4.Text = "BEFORE DELIVERY" And ComboBox2.Text = "RURAL AREA" And TextBox2.Text =
"7&9" And ComboBox3.Text = "GOOD" And TextBox3.Text = "18-20" And TextBox4.Text = "1"
Then
            TextBox1.Text = "NO life status"
        ElseIf ComboBox1.Text = "NORMAL BP" And ComboBox5.Text = "NORMAL LOBOR" Or
ComboBox5.Text = "POOR MATERNAL EFFORTS" And ComboBox4.Text = "BEFORE DELIVERY" And
ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text = "DEBUB" Or ComboBox2.Text = "INNER
CITY" Or ComboBox2.Text = "OTHER" And TextBox2.Text = "6&8" And ComboBox3.Text = "GOOD"
And TextBox3.Text = "18-20" And TextBox4.Text = "1" Then
            TextBox1.Text = "Normal life Status"
        ElseIf ComboBox1.Text = "NORMAL BP" And ComboBox5.Text = "BRADY CARDIA" Or
ComboBox5.Text = "POOR MATERNAL EFFORTS" And ComboBox4.Text = "BEFORE DELIVERY" And
ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text = "DEBUB" Or ComboBox2.Text = "INNER
CITY" Or ComboBox2.Text = "OTHER" And TextBox2.Text = "7&9" Or TextBox2.Text = "8&8" Or
```

```
TextBox2.Text = "8&9" And ComboBox3.Text = "GOOD" And TextBox3.Text = "18-20" And
TextBox4.Text = "1" Then
            TextBox1.Text = "Risky life Status"
        ElseIf ComboBox1.Text = "HIGH BP" And ComboBox5.Text = "NORMAL LOBOR" Or
ComboBox5.Text = "POOR MATERNAL EFFORTS" And ComboBox4.Text = "AFTER DELIVERY" And
ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text = "DEBUB" Or ComboBox2.Text = "INNER
CITY" Or ComboBox2.Text = "OTHER" And TextBox2.Text = "0" And ComboBox3.Text = "GOOD"
And TextBox3.Text = "18-20" And TextBox4.Text = "1" Then
            TextBox1.Text = "Normal life Status"
        ElseIf ComboBox1.Text = "HIGH BP" And ComboBox5.Text = "NORMAL LOBOR" Or
ComboBox5.Text = "POOR MATERNAL EFFORTS" And ComboBox4.Text = "BEFORE DELIVERY" And
ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text = "DEBUB" Or ComboBox2.Text = "INNER
CITY" Or ComboBox2.Text = "OTHER" And TextBox2.Text = "0" And ComboBox3.Text = "GOOD"
And TextBox3.Text = "27-43" And TextBox4.Text = "1" Then
            TextBox1.Text = "Risky life Status"
        ElseIf ComboBox1.Text = "HIGH BP" And ComboBox5.Text = "POOR MATERNAL EFFORTS"
And ComboBox4.Text = "BEFORE DELIVERY" And ComboBox2.Text = "RURAL AREA" And
TextBox2.Text = "0" And ComboBox3.Text = "GOOD" And TextBox3.Text = "18-23" And
TextBox4.Text = "1-34" Then
            TextBox1.Text = "Risky life Status"
        ElseIf ComboBox1.Text = "NORMAL BP" And ComboBox5.Text = "NORMAL LOBOR" Or
ComboBox5.Text = "POOR MATERNAL EFFORTS" And ComboBox4.Text = "BEFORE DELIVERY" And
ComboBox2.Text = "INNER CITY" And TextBox2.Text = "8&9" And ComboBox3.Text = "GOOD" And
TextBox3.Text = "23-43" And TextBox4.Text = "1-4" Then
            TextBox1.Text = "Normal life Status"
        ElseIf ComboBox1.Text = "HIGH BP" And ComboBox5.Text = "POOR MATERNAL EFFORTS"
And ComboBox4.Text = "AFTER DELIVERY" And ComboBox2.Text = "RURAL AREA" Or
ComboBox2.Text = "DEBUB" And TextBox2.Text = "9&6" And ComboBox3.Text = "GOOD" And
TextBox3.Text = "30-43" And TextBox4.Text = "6-34" Then
            TextBox1.Text = "Risky life Status"
        ElseIf ComboBox1.Text = "NORMAL BP" And ComboBox5.Text = "BRADY CARDIA" And
ComboBox4.Text = "BEFORE DELIVERY" And ComboBox2.Text = "RURAL AREA" And TextBox2.Text =
"6&9" Or TextBox2.Text = "7&8" Or TextBox2.Text = "8&8" Or TextBox2.Text = "8&9" And
ComboBox3.Text = "GOOD" And TextBox3.Text = "23-43" And TextBox4.Text = "6-23" Then
            TextBox1.Text = "Risky life Status"
        ElseIf ComboBox1.Text = "MEIUDM BP" And ComboBox5.Text = "NORMAL LOBOR" Or
ComboBox5.Text = "POOR MATERNAL EFFORTS" And ComboBox5.Text = "AFTER DELIVERY" And
ComboBox2.Text = "INNER CITY" And TextBox2.Text = "6&8" And ComboBox3.Text = "GOOD" And
TextBox3.Text = "17-43" And TextBox4.Text = "1-34" Then
            TextBox1.Text = "Normal life Status"
        ElseIf ComboBox1.Text = "HIGH BP" And ComboBox5.Text = "POOR MATERNAL EFFORTS"
And ComboBox4.Text = "AFER DELIVERY" And ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text
= "DEBUB" Or ComboBox2.Text = "INNER CITY" Or ComboBox2.Text = "OTHER" And TextBox2.Text
= "0" And ComboBox3.Text = "GOOD" And TextBox3.Text = "25-43" And TextBox4.Text = "1-34"
Then
            TextBox1.Text = "Normal life Status"
        ElseIf ComboBox1.Text = "HIGH BP" And ComboBox5.Text = "NORMAL LOBOR" And
ComboBox4.Text = "AFER DELIVERY" And ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text =
"INNER CITY" And TextBox2.Text = "6&8" Or TextBox2.Text = "7&9" Or TextBox2.Text = "7&8"
Or TextBox2.Text = "8&8" And ComboBox3.Text = "GOOD" And TextBox3.Text = "17-23" And
TextBox4.Text = "2" Then
            TextBox1.Text = "Normal life Status"
        ElseIf ComboBox1.Text = "HIGH BP" And ComboBox5.Text = "NORMAL LOBOR" And
ComboBox4.Text = "AFER DELIVERY" And ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text =
"INNER CITY" And TextBox2.Text = "6&8" Or TextBox2.Text = "7&9" Or TextBox2.Text = "7&8"
Or TextBox2.Text = "8&8" And ComboBox3.Text = "GOOD" And TextBox3.Text = "17-23" And
TextBox4.Text = "2" Then
            TextBox1.Text = "Normal life Status"
```

```
        ElseIf ComboBox1.Text = "MEDIUM BP" And ComboBox5.Text = "POOR MATERNAL EFFORTS"
And ComboBox4.Text = "AFER DELIVERY" And ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text
= "INNER CITY" And TextBox2.Text = "6&8" Or TextBox2.Text = "7&9" Or TextBox2.Text =
"7&8" Or TextBox2.Text = "8&8" And ComboBox3.Text = "GOOD" And TextBox3.Text = "17-23"
And TextBox4.Text = "2" Then
            TextBox1.Text = "Normal life Status"
        ElseIf ComboBox1.Text = "HIGH BP" And ComboBox5.Text = "BRADY CARDIA" And
ComboBox4.Text = "AFER DELIVERY" And ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text =
"INNER CITY" And TextBox2.Text = "6&8" Or TextBox2.Text = "7&9" Or TextBox2.Text = "7&8"
Or TextBox2.Text = "8&8" And ComboBox3.Text = "GOOD" And TextBox3.Text = "17-23" And
TextBox4.Text = "2" Then
            TextBox1.Text = "Risk life status "
        ElseIf ComboBox1.Text = "MEDIUM BP" And ComboBox5.Text = "POOR MATERNAL EFFORTS"
And ComboBox4.Text = "AFER DELIVERY" And ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text
= "INNER CITY" And TextBox2.Text = "6&8" Or TextBox2.Text = "7&9" Or TextBox2.Text =
"7&8" Or TextBox2.Text = "8&8" And ComboBox3.Text = "GOOD" And TextBox3.Text = "17-23"
And TextBox4.Text = "2" Then
            TextBox1.Text = "Normal life status "
        ElseIf ComboBox1.Text = "MEDIUM BP" And ComboBox5.Text = "BRADY CARDIA" And
ComboBox4.Text = "AFER DELIVERY" And ComboBox2.Text = "RURAL AREA" Or ComboBox2.Text =
"INNER CITY" And TextBox2.Text = "6&8" Or TextBox2.Text = "7&9" Or TextBox2.Text = "8&8"
Or TextBox2.Text = "8&8" And ComboBox3.Text = "NOT GOOD" And TextBox3.Text = "17-23" And
TextBox4.Text = "2" Then
            TextBox1.Text = "Risky life status "
        Else
            TextBox1.Text = "Un Predictable"
        End If

    End Sub

Private Sub Label2_Click(sender As Object, e As EventArgs) Handles L1.Click
        TextBox1.Text = ""
        TextBox2.Text = ""
        TextBox3.Text = ""
        TextBox4.Text = ""
        ComboBox2.Text = ""
        ComboBox3.Text = ""
        ComboBox5.Text = ""
        ComboBox1.Text = ""
        STATUS.Text = ""
    End Sub

Private Sub ListBox1_SelectedIndexChanged(sender As Object, e As EventArgs)
        TextBox1.Text = ""
        TextBox1.Text = ""
        TextBox2.Text = ""
        TextBox3.Text = ""
        TextBox4.Text = ""
        ComboBox2.Text = ""
        ComboBox3.Text = ""
        ComboBox5.Text = ""
        ComboBox1.Text = ""
        STATUS.Text = ""
    End Sub

Private Sub STATUS_Click(sender As Object, e As EventArgs) Handles STATUS.Click
        STATUS.Text = ""
    End Sub
```