

# *i*Birth: An intelligent web application that recommends childbirth delivery mode

By

Uzapi Hange

Reg. No: 14101345

BSc (Applied Business Computing) (University of Sunderland)

Department of Computer Science and Software Engineering,

College of Information and Communications Technology,

Botswana International University of Science and Technology

uzapi.hange@studentmail.biust.ac.bw, (+267) 765 715 21

A Research Proposal Submitted to the College of ICT for the Study Leading to a Dissertation in Partial Fulfilment of the Requirements for the Award of the Master of Science in Computer Science of BIUST

Supervisor(s): Dr Rajalakshmi Selvaraj

Department of Information systems

College of ICT, BIUST

[**Selvarajr@biust.ac.bw**](mailto:Selvarajr@biust.ac.bw)**,** [**+267-72445426**](tel:%2B267-72445426)

Signature \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

May, 2015

# Abstract

Maternal health is one aspect that is important in the Botswana health system. If not given the attention it deserves it can lead to serious complications or maternal death which is a worldwide concern. Childbirth type prediction is a facet of maternal health and its importance cannot be overlooked. Expectant mothers and midwifes need to know their delivery mode well in time to plan the childbirth in a manner that will make it a positive experience. The traditional method that involves human efforts is still being used in Botswana. This method has a lot of setbacks that are not limited to emergency C-Sections, erroneous classifications and time consuming. With the available computer science technologies like decision trees and visualisation that make decision trees interactive and more interpretable; there is a need to develop an application that classifies birth type in an interactive manner. Consequently the study proposed a web application (*i*Birth) that will support midwifes and medical staff birth type classification decision making. *i*Birth’s goal is to develop a web application that classifies childbirth type into natural birth or C-Section using a visualised C4.5 decision tree that makes classification interactive, via a zoomable interface that allows selection, filtering of data and viewing results in various visualisation forms such as 2D and 3D.

To implement the application programming languages and technologies that include Visual Basic, HTM, CSS, JavaScript will be used. These will be implemented on ASP.NET or Visual studio platforms. The interactive visualisation will be incorporated using D3.Js (a JavaScript library that visualises web applications). An experiment will be conducted to test the mean absolute error, and convergence time of C4.5 and the interactivity of *i*Birth.

# Key words

Childbirth, data mining, maternal health, classification, prediction, visual data mining, information visualization, interactive, C4.5, feature selection, gain ratio, decision tree, decision support system, midwife, classifier

Table of Contents

[*i*Birth: An intelligent web application that recommends childbirth delivery mode 1](#_Toc420193417)

[Abstract 2](#_Toc420193418)

[Key words 2](#_Toc420193419)

[List of abbreviations 4](#_Toc420193420)

[Chapter 1 introduction 5](#_Toc420193421)

[**Introduction to the problem** 5](#_Toc420193422)

[Background 7](#_Toc420193423)

[Problem statement 8](#_Toc420193424)

[General objective 9](#_Toc420193425)

[Specific objectives 9](#_Toc420193426)

[Hypothesis/research questions 9](#_Toc420193427)

[Expected outcomes 9](#_Toc420193428)

[Justification of the study 10](#_Toc420193429)

[Chapter 2 literature review 11](#_Toc420193430)

[Proposal structure 17](#_Toc420193431)

[Chapter 3 methodology 18](#_Toc420193432)

[References 20](#_Toc420193433)

[Appendix A Schedule of activities 22](#_Toc420193434)

[Appendix B indicative research budget 23](#_Toc420193435)

[Appendix C Proposed web application architecture 24](#_Toc420193436)

# List of abbreviations

**C-SECTION –**Caesarean section

**DSS –**Decision support system (s)

**HTML-**Hypertext markup language

**CSS-**Cascading style sheets

**UML-**Unified Modeling Language

# Chapter 1 introduction

Introduction to the problem

Childbirth is an awesome experience but it is not always the case for all expectant mothers. It is often associated with erratic life threatening maternal complications especially in least developed or developing countries. Some disorders such as problems with the placenta, the baby not getting oxygen, or the baby being in an abnormal position; they can put an expectant mother or baby at a higher risk for problems. These problems are not limited to maternal death, pelvic pain, disability; [3] therefore the most relevant child delivery method becomes vital.

Child delivery options are natural birth or Cesarean section (C-Section). By default natural birth is deemed as a safer mode of delivery but since birth varies with women, this study concludes that there is no one universal correct delivery type. Making a decision about the delivery mode that will ensure a less risky and positive birth experience for both the mother and child is ultimate, and in some instances a C-Section has possible benefits that need to be measured against risks. Each method is influenced by myriad factors and has complications that are unique to it [3], and hence child delivery type prediction becomes a significant classification problem. A paper by [1] describes classification as supervised learning method that “maps data into predefined classes”. It is called supervised learning since the classes are known before the actual classification transpires. The classes are assigned based on the data attribute values. The classes are described by often on the basis of the properties of data that is already known to belong to a certain class.

In Botswana, childbirth type classification is being practiced manually by midwifes and medical staff. Every expectant mother is assumed that they are going to deliver naturally until complications arise [7]. As a result a lot of emergency C-Sections have been incurred, but it is necessary to plan C-Section well beforehand to avert risks that come with it such as permanent physical damage to the baby or mother. The conservative way of classification portrays various insufficiencies in that birth type classification job requires a midwife to access an array of factors and a combination of these factors before making a decision on whether the mother will deliver through C-Section or natural birth. This task becomes almost unattainable since the choice of the delivery mode is solely based on the midwife’s experience, handling too much data manually can get tedious, difficult and time consuming and inexperienced midwifes are bound to make erroneous predictions. These drawbacks obstruct the growth of Botswana maternal health care. Conversely an immense volume of medical and maternal data is accumulated through the manual prediction method which offers a great opportunity to counter the problems of the current system. Patterns and trends from this data could be generated and used for future classification, and data mining has the potential to complete this undertaking. Some advancement have been made in applying data mining techniques to medical health DSS and childbirth classification DSS. The techniques include neural networks, logistic regression, support vector machines and decision trees. These algorithms decrease the time spent for processing symptoms and producing diagnoses and classification, making them more precise at the same time [5].

However decision they sometimes produce large trees and classification rules that can be difficult to comprehend and interact with. They do not allow for the user to understand why a certain decision was made and consequently choose to adopt or discard the outcome of the classifier. This inadequacy demands an automated classifier that accurately performs the classification exercise and allows for visualization and let users interact with the classifier results. Fusing the available decision trees like C4.5 and Interactive visualisation technologies can solve the manual childbirth classification problem. Information visualisation research offers many techniques that help decision support systems users to explore, understand and analyse data in a coherent manner. It achieves this by exhibiting data in a graphical format and using interactive interfaces so as to best express complicated concepts in a comprehensive way, to simplify interpretation, causal relationships and dependencies. Information visualization has been successfully used in data mining, scientific research, financial data analysis, market analysis, and manufacture control studies [2]. And therefore this study sees the necessity to implement information visualisation in childbirth type classification. A web application that uses C4.5 decision tree to carry out childbirth classification and an interactive interface that visualises the C4.5 results is proposed. The interface intends to offer functionalities such as selection of area of interest, filtering, zooming and displaying the results in multiple formats.

## Background

Delivery type classification is vital for preparing for childbirth while there is still time. Different medical DSS, maternal DSS and visualisation systems are discussed below;

In Botswana, childbirth type classification is manual. It is assumed that an expectant mother is going to deliver naturally until complications arise [4]. This has led to a lot of emergency C-Sections, classifying childbirth type beforehand is necessary to avoid risks that come with it.

In their study [6] presented a conventional logistic regression, a statistical model, to develop a model for predicting the success of a test of vaginal birth after Cesarean section based on ultrasound measurements of C-Section scar features. It achieved an accuracy rate of 61%. Nonetheless they produce results in traditional formats that are not interactive.

A birth prediction classification model (BPC) was proposed by [7] based on three classifiers; J48 decision tree, apriori association rule algorithm, and neural networks. It was able to classify the births into natural and C-Section with an average accuracy, precision and recall of 80%, 85% and 84% respectively. Association rule mining rule was used to draw attention to the imperative medical factors that are linked with C-Section births**.** Decision Tree (J48) classification rate was 80% and artificial Neural Network 82%. The classification rules and trees produced lacked visualisation to make them comprehensive.

In China, an intelligent maternal and child health information system was initiated in an attempt to solve the problem of shortage of medical resources. It combined cloud computing and data mining technologies. The system increased medical data processing speed and offered a variety of healthcare service alternatives [8]. However like other decision support systems it lacked visualisation of data mining results, but for data mining to be complete, the data mining results have to be well interpretable and this is accomplished by visualisation.

To overcome setbacks of these systems, research by [2] suggests adoption of information visualisation research in visualising DSS results. Various systems that use information visualisation are discussed below;

VISCARETRAILS system was developed by [12] to allow users to explore electronic health data to understand patterns, problems and opportunities in clinical settings. It uses timed word tree visualizations to summarize event paths in relation to a certain root event. It is only useful when analyzing records where there is a discrete start and end event.

Research by [9] developed a system named PaintingClass. It involves humans in the decision tree construction and supports interaction as well[2].

Another system named MGV: A System for Visualizing Massive Multidigraphs allows for flexible interactivity with graphs. It has functionalities that include a zoomable interface, filter, and incorporating a number of visualization techniques like 2D and 3D.

After thrashing out the above systems we conclude that decision support systems have been used in maternal and other medical data to support maternal health operations such as diagnostics and prognosis of diseases. These systems decrease the time spent for making diagnoses and classification, and assist in making medical diagnosis a lot more accurate. However most of the systems do not offer interactive classification features. Botswana midwifes lack decision support systems that aid childbirth delivery mode classification. And therefore merging computer science technologies to implement systems that perform classification is vital. *I*Birth seeks to merge the pluses of MGV with those of a C4.5 decision tree to make C4.5 interactive. The user should be able to use the web application to classify births and be able to understand the classifier results and interact with them. The web application will provide a zoomable interface, it will also allow selection of area of interest, filtering and view of the C4.5 in different visualizations such as 2D and 3D. This is what makes this study unique.

## Problem statement

Botswana maternal healthcare lacks systems that support classification of childbirth delivery mode. However other available childbirth type classification DSS that the study attempts to adopt produce outcomes in conventional formats (decision trees or classification rules) that are tedious and difficult to interpret. The problem that the study will address is lack of use of visualization approaches in childbirth type classification DSS to make the knowledge discovery process interactive.

## General objective

To develop a web application, that will recommend patient specific child delivery mode based on a C4.5 classifier and allow for interaction with the classifier outcome.

## Specific objectives

1. To collect and organise maternal data.

2. To design the web application including the graphical user interface and visualisation.

3. To develop the web application including the graphical user interface, C4.5 and feature selection algorithm and interactive visualisation

4. Train C4.5 childbirth type classification model

5. Test C4.5 childbirth type classification model in terms of mean absolute error and convergence time. And interactivity using the following matrices; zoomable interface, selection of area of interest, filtering and viewing C4.5 in different visualizations such as 2D and 3D)

## Hypothesis/research questions

* What attributes affect each childbirth type?
* What are the requirements for implementing a web based childbirth type classification model?
* How can information visualisation be incorporated into the web based childbirth type classification model?
* How does information visualization improve interaction, will it make the web application that classifies childbirth type interactive?
* How will C4.5 and information visualisation improve childbirth type classification?

## Expected outcomes

The outcome of the research is a visualised C4.5 decision tree and a web application for midwifes that will predict patient specific child delivery mode based on the visualised C4.5 decision tree. A zoomable interactive interface will allow selection of area of interest, filtering and view of the C4.5 in different visualizations such as 2D and 3D.

## Justification of the study

The study aims at developing a web application that will assist midwifes to predict child delivery type. This procedure is currently done manually by midwifes in Botswana. Due to shortage of midwifes in Botswana, sometimes non-midwife medical staff is expected to perform maternal health duties [4], but experience is vital to perform a task of such a massive magnitude. As a result of the difficulties encountered through manual childbirth type classification, the demand for a novel system that handles the huge amount of maternal data, in a fast and comprehensive manner escalates. The web application will classify childbirth type in advance to avoid complications that accompany childbirth classification at a very late stage. The application is primarily developed for midwifes but it will also be of great help to other inexperienced medical staff.

The application will incorporate data mining and information visualization techniques to contribute to the computer science body of knowledge, an improved C4.5 decision tree classification model based on an interactive interface. The new application will visualize the results to make them more interpretable for the midwife. Midwifes will be able select of area of interest, filter data and view of the C4.5 in different visualizations such as 2D and 3D. An undertaking of this kind will be new in Botswana and so this study is worth pursuing.

# Chapter 2 literature review

This chapter discusses the various current maternal and other medical related DSS that are in place. Systems that employ data mining algorithm and information visualization techniques are also plainly thrashed out in this section.

In Botswana the birth prediction system is done manually and is solely dependent on screening and physical examination for health conditions; blood pressure, weight, urine test, and scan results to name a few. And the final decision relies on the midwife’s knowledge and experience. However, some studies reveal that with the advent of technology, there are various maternal health DSS that have been introduced to assist the medical staff to make important decisions. A few DSS are scrutinized below;

**MATERNAL AND MEDICAL DECISION SUPPORT SYSTEMS**

**Automated Diagnosis and Cause Analysis of Cesarean Section Using Machine Learning Techniques**

Research by [7] Developed a Birth Prediction Classification (BPC) model based on three classifiers; J48 decision tree, apriori association rule algorithm, and neural networks. It successfully managed to classify the births into natural and C-Section with an average accuracy, precision and recall of 80%, 85% and 84% respectively. Association rule mining rule set could draw attention to the imperative medical factors that are linked with C-Section births**.** Decision Tree (J48) classification rate was 80% and artificial Neural Network 82%. The author concluded that the results are satisfactory but can be improved by identifying additional factors that influence birth type and that the model could be more useful if accessible online.

**Predicting successful vaginal birth after Cesarean section using a model based on Cesarean scar features examined by transvaginal sonography (ultrasound)**

A study by [6] used a conventional mathematical approach (logistic regression) to develop a model for predicting the success of a test of vaginal birth after Cesarean section based on ultrasound measurements of C-Section scar features, demographic variables and previous obstetric history. It achieved a prediction rate of 61%. The use of conventional statistical methods make research insufficient, they are time consuming, and produce results in traditional formats that lack interaction functionalities.

**An Intelligent Information System for Maternal and Child Health Care**

In China, an intelligent maternal and child health information system was initiated in an attempt to solve the problem of shortage of medical resources. It is based on cloud computing and data mining technologies. The system increased medical data processing speed and offered a variety of healthcare service alternatives [8]. However it has a shortfall in that it lacks visualisation of data mining results, but for data mining to be complete, its results have to be well interpretable and this is accomplished by visualisation. The system is exhibited by figure 2 below;



Figure 2

**Application of data mining methods for and techniques for diabetes classification**

In their research work, [1] proposed a system that applied different data mining techniques to classify Diabetes data and forecast the possibility of a patient having Diabetes or not. Among the Classification Algorithms was the RND TREE that provided an accuracy rate of 100% but the algorithm experienced data over fitting problems and the rule set obtained was massive. On the other hand, C4.5’s accuracy rate was approximately 91% classification. C4.5 is mainly used for most of the medical applications, for this reason it was used for the classification. The study concluded that in future C4.5 needs to be improved to achieve a greater classification rate. A survey by [2] states that information visualisation can handle noise and inhomogeneous and makes data exploration fast and yields better outcomes when data mining algorithms fail and so there is room to improve C4.5 using information visualisation.

**VISCARETRAILS system**

VISCARETRAILS system was developed by [12] to allow users to explore electronic health data to understand patterns, problems and opportunities in clinical settings. It uses timed word tree visualizations to summarize event paths in relation to a certain root event. The word trees are generated using a drag-and-drop user interface. This is useful when analyzing records where there is a discrete start and end event such as cancer records. Nevertheless the challenge is in trying to optimizing it to carry out both diagnostic testing strategies and bundled events that are associated with differences in survival. It is also not combined with any algorithmic method to aid classification; it only supports record analysis which is not adequate.

**WHY VISUALISATION?**

Though high accuracy rates were achieved by the algorithms mentioned above, their generated rules were massive and could hinder interpretation of the results. And this could render the algorithms useless since the main objective of data mining is to extract hidden knowledge and be able to understand and utilize them in making important decisions. Though VISCARETRAIL system supported visualisation it was only for the purpose of record viewing. Visual data mining was birthed as a result of the limitations of the algorithms discussed above. Several visualization approaches are currently being used to aid better interpretation of results and human computer interaction. The approaches are critiqued below;

**PaintingClass: Interactive Construction, Visualization and Exploration of Decision Trees**

PaintingClass, is a system that builds,visualizes and explore decision trees. In PaintingClass, The user is enabled to interactivelyedit predictions of multi-dimensional data and paint aregion to build a decision tree. PaintingClass’s effectiveness was experimentally verifiedby applying it to classification of actual data with up to 19 dimensions, and comparing itsperformance to that of well-known algorithmic and visual classificationmethods [9]. The accuracy of the algorithms is as follows;

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dataset** | **CBA** | **C4.5** | **FID** | **Fuzzy** | **PaintingClass** |
| Australian | 85 | 82.6 | 58.0 | 88.9 | 84.7 |
| Adult | 84.2 | 85.4 | 23.6 | 85.9 | 85.1 |
| Diabetes | 74.4 | 73.8 | 62.0 | 77.6 | 74.6 |

The writer concluded that, with additional development like incorporatingalgorithmic techniques, PaintingClass can yield better results in terms of accuracy.

**Descartes system**

Descartes system [13] (used to be called IRIS) was developed to provide visual exploration of spatial data like demographic, cultural, economical, geographic information or locations such as countries. It made use of C4.5 and offered automated depiction of data on maps and interactive manipulation services. Descartes looks at the attributes to be analysed and the correlations between the attributes and chooses the most apt presentation method. The cartographic knowledge of Descartes allows non-cartographers to receive proper presentations of their data, and the automation of map construction helps the users to save valuable time that can better be used for data analysis and problem-solving. Different types of icons, diagrams, colored faces, and maps are used for depicting data within their spatial frame of reference. This system, however, does not support the visualization of complex information structures like abstract node link graphs or hierarchies.

**MGV: A System for Visualizing Massive Multidigraphs**

A system by [14] provides an “integrated visualization and exploration system for massive multidigraph navigation. It adheres to the Visual Information-Seeking Mantra: overview first, zoom and filter, then details on demand. MGV's only assumption is that the vertex set of the underlying digraph corresponds to the set of leaves of a predetermined tree T. MGV builds an out-of-core graph hierarchy and provides mechanisms to plug in arbitrary visual representations for each graph hierarchy slice. Navigation from one level to another of the hierarchy corresponds to the implementation of a drill-down interface. In order to provide the user with navigation control and interactive response, MGV incorporates a number of visualization techniques like interactive pixel-oriented 2D and 3D maps, statistical displays, color maps, multilinked views, and a zoomable label based interface. This makes the association of geographic information and graph data very natural. To automate the creation of the vertex set hierarchy for MGV, we use the notion of graph sketches. They can be thought of as visual indices that guide the navigation of a multigraph too large to fit on the available display”.

**Conclusion of literature review**

All the reviewed systems have pros and cons. Most if not all the clinical decision support systems do not support information visualization. However, merging them with interactive methods could improve their performance and aid human interaction since it is not enough to only present the data in graphical format. Understanding the classifier results and being able to navigate the results in an interactive way is essential and this task is completed by information visualisation. It combines the flexibility and overall knowledge of the human with the large storage size and computational power of modern computer systems. In essence it presents data in visual form, allow humans to get insight into patterns, make conclusions and interact with data directly. Visual data exploration has proven to be beneficial especially when only a handful is known about the data. Some of the merits of visualisation are as follows; it can handle noise and inhomogeneous data, it requires no prior understanding of complex statistical algorithms, thus making information visualisation research profound.

It is very vivid in the literature that the systems that use visualisation allow for flexibility in terms of navigation and user interaction. *i*Birth will benchmark on the strong points of two systems from the literature; an Automated Diagnosis and Cause Analysis of Cesarean Section Using Machine Learning Techniques research and MGV system.

The new system will employ C4.5 decision tree which is known as J48 in WEKA to build a childbirth type classification model. The chief reason for this is because C.5 has been used in many medical areas with high accuracy rates. It also performs feature selection using Gain ratio to select attributes that are relevant and remove the ones that are irrelevant for training. This will help manage the numerous features that are involved in determining childbirth type. On the other hand the functionalities that will be borrowed from the MGV system are the following, zoomable interface, filter, and incorporating a number of visualization techniques like 2D and 3D. Another feature that will be added is selection of point of interest. The merits of these two systems will provide an interactive and flexible childbirth type classification web application.

## Proposal structure

The proposal is organized as follows; title page, abstract, table of contents and list of abbreviations. Chapter one that consists of introduction, Background, problem statement, general and specific objectives, hypothesis/research questions, expected outcomes, justification of the study, and proposal structure. Chapter two reviews the current literature. Chapter three focuses on outlining the methodology. It is followed by appendix A which has the gantt chart and appendix B contains the indicative research budget and finally the reference list.

# Chapter 3 methodology

Figure 2 illustrates the steps that will be taken to achieve the goal and objectives of the research. Each objective and how it is going to be achieved is explained below;

**To collect and organize maternal data**

Maternal data will be collected from the ministry of health in Gaborone. This data will be used for training and testing our model. Semi-structured interviews with medical staff will also be conducted so as to explore the most common factors that influence birth type and the risks and complications associated with each type. Published and unpublished documents on maternal health and child care will also be reviewed.

Not all the collected data will be needed for the study and it might compromise the quality of the study as it is regarded as noise. And therefore collected data will be preprocessed and organized, this will include data cleaning, putting into the desired format, and verifying the quality of the data. When data organization is done, the data will be separated into training and test set. It will then be loaded into a database.

**To design the web application including the graphical user interface and visualisation**

The interface that the users are going to make use of to interact with the C4.5 algorithm and the results, the look and feel of the web application, and how the visualisation is going to be implemented will be all designed using various tools such UML Diagrams, Screen flow diagrams, use case diagrams and wireframes.

**To develop the graphical user interface, C4.5 and feature selection algorithm and interactive visualisation**

The development of the user interface will be carried out using HTML and CSS. The C4.5 and gain ratio will be coded using visual basic programming language on either web matrix platform or ASP.NET or Visual studio. The interactive visualisation will be developed using D3.Js

**Train C4.5 childbirth type classification model**

The training of the developed C4.5 classification model will take place in the ASP.NET or VB.net framework.

**Test C4.5 childbirth type classification model**

Gain ratio feature selection method will be applied the training set. To select attributes that are relevant and remove those that are irrelevant for training. After feature selection, the training set will be loaded on to the developed model to train it. This is where patterns and knowledge is going to be extracted. The mean absolute error (the difference between the data that was correctly classified by the training data and the one that was correctly classified by the test data), time taken to converge and interactivity (zoomable interface, selection of area of interest, filtering and view of the C4.5 in different visualizations such as 2D and 3D) will be tested through an experiment.

The methodology is summarized below;

|  |  |  |
| --- | --- | --- |
|  | Specific Objective | Methods |
| 1 | To collect and organize maternal data | Interviews and document review |
| 2 | To design the web application including the graphical user interface and visualisation. | UML Diagrams, Screen flow diagrams, use case diagrams and wireframes |
| 3 | To develop the graphical user interface, C4.5 and feature selection algorithm and interactive visualisation | ASP.NET/VB.Net, Web matrix, Microsoft SQL server, HTML, CSS, JavaScript, Visual Basic, D3.Js |
| 4 | Train C4.5 childbirth classification model | VB.Net |
| 5 | Test C4.5 childbirth classification model | Mean absolute error, convergence time, interactivity (zoomable interface, selection of area of interest, filtering and viewing C4.5 in different visualizations such as 2D and 3D) |

# References

[1] K. Rajesh and V. Sangeetha, “Application of Data Mining Methods and Techniques for Diabetes Diagnosis,” *Int. J. Eng. Innov. Technol.*, vol. 2, no. 3, pp. 224–229, 2012.

[2] M. Khan and S. S. Khan, “Data and Information Visualization Methods , and Interactive Mechanisms : A Survey,” vol. 34, no. 1, pp. 1–14, 2011.

[3] A. B. E. Review, “Vaginal or Cesarean Birth : What Is at Stake for Women and Babies ?”

[4] T. T. Fako, N. Forcheh, and E. Ncube, “Prospects of safe motherhood in Botswana : midwifery training and nurses ’ ability to complete the Botswana obstetric record,” vol. 58, pp. 1109–1120, 2004.

[5] H. D. Masethe and M. A. Masethe, “Prediction of Heart Disease using Classification Algorithms,” 2014, vol. II, pp. 22–24.

[6] O. Naji, L. Wynants, A. Smith, Y. Abdallah, C. Stalder, and A. Sayasneh, “Predicting successful vaginal birth after Cesarean section using a model based on Cesarean scar features examined by transvaginal sonography,” no. December 2012. pp. 672–678, 2013.

[7] A. Sana, S. Razzaq, and J. Ferzund, “Automated Diagnosis and Cause Analysis of Cesarean Section Using Machine Learning Techniques,” *Int. J. Mach. Learn. Comput.*, vol. 2, no. 5, pp. 677–680, 2012.

[8] L. Chen, J. Yang, Q. Wang, and M. Wang, “An Intelligent Information System for Maternal and Child Health Care,” in *2011 35th IEEE Annual Computer Software and Applications Conference Workshops*, 2011, pp. 441–445.

[9] S. T. Teoh, “PaintingClass : Interactive Construction , Visualization and Exploration of Decision Trees.” 2003.

[10] M. B. Nejad, E. B. Nejad, and A. Karami, “Using Data Mining Techniques to Increase Efficiency of Customer Relationship Management Process,” *Res. J. Appl. Sci. Eng. Technol.*, vol. 4, no. 23, pp. 5010–5015, 2012.

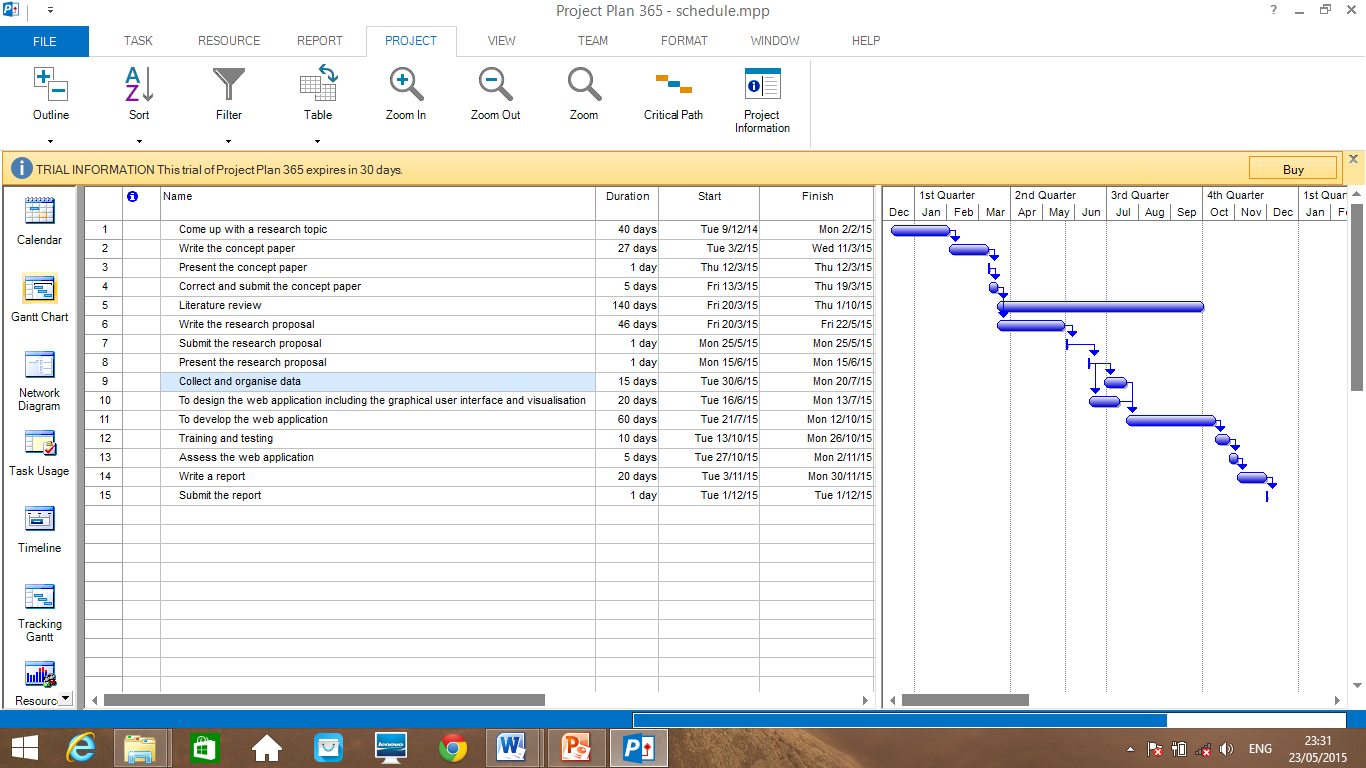
[11] M. Kreuseler and H. Schumann, “A Flexible Approach for Visual Data Mining,” vol. 8, no. 1, pp. 39–51, 2002.

[12] L. Lins, M. Heilbrun, J. Freire, and C. Silva, “V IS C ARE T RAILS : Visualizing Trails in the Electronic Health Record with Timed Word Trees , a Pancreas Cancer Use Case,” pp. 1–4, 2014.

[13] G. L. Andrienko and N. V Andrienko, “Data Mining with C4 . 5 and Interactive Cartographic Visualization,” pp. 6–9.

[14] J. Abello and J. Korn, “MGV : A System for Visualizing Massive Multidigraphs,” *IEEE Trans. Vis. Comput. Graph.*, vol. 8, no. 1, pp. 21–38, 2002.

# Appendix A Schedule of activities



# 

# Appendix B indicative research budget

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **BOTSWANA INTERNATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY** | | | | | | | | |
| **Research title:** *i*Birth: An intelligent web application that recommends child delivery mode | | | | | | | | |
| **INDICATIVE RESEARCH BUDGET** | | | | | | | | |
| **S/N** | **Activity description** | **Input item** | **No of units** | **Unit cost (BWP)** | **Total cost (BWP)** | | **Time frame** | **Location** |
| **Student** | **School** |
| 1 | Data collection | Transport | Palapye - Gaborone | P400 | P0.00 | P400 | 30 June-20 July | Gaborone |
| Food | 10 days | P200 | P.00 | P2 000 |
|  | **Sub Total** |  | P6 000 |
| 2 | Development of the application | Laptop | 1 | P7000 |  | P7 000 | 21 July-12 Oct | BIUST |
|  |  | **Sub Total** |  | P7 000 |
|  | | | | **Grand Total** |  | **P13 000** |

# Appendix C Proposed web application architecture

Feature selection (Gain ratio)

Training data

Tested model

Develop the classification model (C4.5)

Visualisation library (Visualises C4.5 results as requested by users via an interface)

Test C4.5

Testing data

 Midwife/medical staff request