JIMMA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF INFORMATION SCIENCE

Developing Knowledge Based Expert System for Road Traffic Accident Control:

The Case of Jimma Town Traffic Office

By

Kejela Tasissa

June, 2017

Jimma, Ethiopia

JIMMA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF INFORMATION SCIENCE

Masters of Information Science Specialized in IKM

DEVELOPING KNOWLEDGE BASED EXPERT SYSTEM FOR ROAD TRAFIC ACCIDENT CONTROL: THE CASE OF JIMMA TOWN

Thesis Submitted to the Department of Information Science, College of Natural Science, Jimma University in Partial Fulfillment of the Requirements for Degree of Masters of Science in Information Science (Information and Knowledge Management)

BY

KEJELA TASISSA

Principal Advisor:-Martha Yifiru (PhD)

Co-Advisor:- Amanuel Ayde (MSc)

June, 2017

Jimma, Ethiopia

JIMMA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF INFORMATION SCIENCE

DEVELOPING KNOWLEDGE BASED EXPERT SYSTEM FOR ROAD TRAFIC ACCIDENT CONTROL: THE CASE OF JIMMA TOWN

BY

KEJELA TASISSA

As members of the board of examining of the MSc thesis open defense examination of the above title, we members of the board (listed below), read and evaluated the thesis and examined the candidate.

Name	**Title**	**Signature**	**Date**
----------	Chair person	----------	----------
Martha Yifru (PhD)	Principal Adviser	----------	----------
Amanuel Ayde (Msc)	Co Adviser	----------	----------
----------	External examiner	----------	----------
----------	Internal examiner	----------	----------

DECLARATION

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

_____

Kejela Tasissa

June, 2017

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

Principal Advisor: Martha Yifru (PhD) …….

Co-Advisor: Amauel Ayde (MSc) ……

June, 2017

DEDICATION

This work is dedicated to my lovely wife Hana Hailu.

Acknowledgement

First and foremost, I would like to give a special gratitude to the Gracious God who provided me everything to finish my courses and lead me to conduct this research. I gratefully acknowledge my Principal adviser Dr. Martha Yifru and co-advisor, Mr. Amanuel Ayde for their commitments and patience for reading every section of this research, their valuable comments, encouragement and guidance from initial to the final level of the research.

I express my warm thanks to Jimma University Department of Information Science staffs, my families and all the friends and colleagues who provided me with the facilities being required and conductive conditions for this research

I would also like to thank Jimma traffic police head of Alazar and Mentina zone for their kindly help and support. Especially Mr. Yasin and Mr. Nasiru this research would not have been successful without your help and encouragement.

I would like to acknowledge some of Jimma traffic police station members for providing important information for my research work and willing of the traffic police members to help me anywhere.

Finally, I would like to thank my beloved wife for her moral support and encouragements throughout my study and research.

Table of Contents

CONTENTS **PAGE**

List of tables

List of figures

List of Abbreviations and Acronyms

AI	Artificial Intelligence
ARFF	Attribute-Relation File Format
CSS	Cascading Sheet
CSV	Comma Delimited value File
DM	Data Mining
DOT	Department of Transportation
ES	Expert System
FRED	Freeway Real time Expert System Demonstration
HCM	Highway Capacity Manual
HTML	Hyper Text Markup Language
IE	Inference Engine
KADS	Knowledge Acquisition and Documentation Structuring
KDD	Knowledge Discovery in Database
KBES	Knowledge Based Expert System
KBS	Knowledge-Based System
KE	Knowledge Engineer
KR	Knowledge Representation
LIIS	LPR Intelligent Information System
LPR	Logic of Plausible Reasoning
OECD	Organization of Economic Co-operation and Development
RTAC	Road Traffic Accident Control
RTACKBES	Road Traffic Accident Control Knowledge Based Expert System
PHP	Hypertext Pre Processor
WEKA	Waikato Environment for Knowledge Analysis
WHO	World Health Organization
WZTS	Work Zone Safety Training System

Abstract

In Ethiopia road traffic accidents are becoming big issue that needs critical solution by government body with traffic police office to save life of the citizens. The purpose of this study is to explore the possibility of developing road traffic accident control knowledge based expert system (RTACKBE) prototype for making decision. RTACKBE system is a system that diagnoses a road traffic accident by learning domain knowledge generated by expert and produce preventive measurement as a solution. The system present mitigating measures and decide potential alternatives solution for alleviation of traffic accidents by considering preventive measurement to save life of citizen. This RTACKBES is scoped with vehicle to vehicle crash and vehicle crash pedestrian. The source of data were from both primary (using such as observation and interview); and secondary data sources like traffic accident report documents for design of this study. To design system prototype researcher used KBS development methodology that involves knowledge acquisition, representation and modeling steps. Moreover, domain knowledge has been acquired from domain expert using data collecting method particularly by interviewing and from data mining techniques to process past traffic accident record. Acquired knowledge was represented by sets of rule determined by PART algorithm for extracted knowledge from data mining. Acquired and represented knowledge was modeled based on decision tree structure resulted by J48 prune algorithm. Finally Web development tools and Waikato Environment for Knowledge Analysis (*WEKA) data mining tools have been used for the design of the system. The performance of system was evaluated and tested using two evaluation performance; one: using test case and two: using user acceptance. As a result of finding the prototype knowledge based expert system is promising and applicable in the domain area (Jimma town) with the user acceptance of 84.17%. The performance of the system was achieved with 84.2% due to perform and process expected task operation. The importance and requirement of the system on domain area was evaluated by domain expert based on simplicity in terms of its efficiency and ease of using was 97.5%. The user accepted RTACKBE system as a result of user acceptance and test case evaluation and performance results are achieved the objectives of this study.*

Chapter One

Introduction

1.1. Back ground of the study

Traffic Agency is responsible government organization for tasks such as road, car traffic accident controlling, road traffic accident data management, etc. Traffic on roads may consist of pedestrian , ridden or herded animals, vehicles, streetcars, buses and other conveyances, either singly or together, while using the public way for purposes of travel. Traffic congestion is a condition on transport networks that occurs as use increases, and is characterized by slower speeds, longer trip times, and increased vehicular queuing. The most common example is the physical use of roads by vehicles.

Road traffic accident (RTA) is an accident that probably happen when vehicle crash people or other vehicle on the road because of different factors. According to global status report (2013) road traffic accident (RTA) is any injury due to crashes originating from, terminating with or involving a vehicle partially or fully on a public road. As a result risk factors that contributed in road traffic accident described on Global status report are mechanical factor, Human factor, driver factor, over speeding increases, reckless driving like use of mobile phones during driving, non-use of helmets, non-use of seat-belts, driver fatigue and sleepiness also contribute to crashes, improper designing of roads and lack of pedestrian pavement are other contributing factors, road factors and Pedestrian factors (Global status report on road safety 2013).

Road traffic accidents are also problems that usually occur in developing country due to the constant mobility of people and goods. With the increase of automobile usage, pedestrians, increase of vehicle in urban as well as city road the number of accidents also increased.

Moreover road traffic accident in Ethiopia is described by different researcher and organization. According to the World Health Organization (2009), traffic accident is "Epidemic of Civilized Societies". It is sad to note that these people just wanted to go home, visit a friend, or go to their work but encountered something terrible along the way. Road transportation has a direct connection with the day to day activities of people, especially in large cities where the distance is too far to cover on foot or by bicycle within a reasonable time (WHO 2009). However, the increase in road transportation has placed a considerable burden on the people's lives. The

fatality of road traffic deaths and injuries is the major one. WHO also asserted that road traffic crashes are the leading causes of death and disablement of people under the age of 44 years next to HIV/AIDS. Road traffic accidents have been growing in Ethiopia overtime, with the exception of the two recent years of 2007/08 and 2009/104. The registered growth is presumed to be significantly under-stated due to the challenge of getting the exact number of accidents occurring throughout the country. WHO rated road traffic accident in Ethiopia (region, zone, cities and urban) the rate of road traffic accidents (RTAs) is very high; because road transport is the major transportation mechanism along with poor road infrastructure, poor traffic laws enforcement and other factors. The Ethiopian traffic control system archives data on various aspects of the traffic system, such as traffic volume, concentration, and vehicle accidents (WHO, 2009).

According to AMC (2015) accident management is the centralized handling of a motorist's claim following a road traffic collision or other damages or mishaps that happen to a vehicle while on or off road. It is a cost-effective intermediary service which assists drivers in getting back on the road quickly and in managing the claims process alone. (Accident Management Center 2015) Whilst it is significantly more cost effective for the innocent motorist, the service costs significantly more as a result - a cost borne by the insurer of the 'at-fault' driver.

Therefore, road traffic accident is attractive area of research which involve complex problem and wanted to be solved. Many international and local researchers investigated on this area to find solution for the accident problem reduction. Most of researchers used data mining techniques to demonstrate on the area where the accident is occurred and injury severity is high, less, loss of property and mortal. Based on the finding and outcome of their research thus researcher recommended their future work as it is important developing effective kind of decision support system in Ethiopia due to reduce accidents more. This system can be such as Expert systems (ES), web based system (WBS), knowledge based system (KBS) and other artificial intelligence technologies such as knowledge based expert system (KBES) which have high potential of contributing in control of road traffic accident.

An expert system (ES) is always meaningful by definition because of the type or field of knowledge that it uses, and its purpose is to apply special knowledge to specific requirements, not simply to retrieve or expose information. It can be defined as a computer system that

emulates the decision-making ability of a human expert. Expert systems are designed to solve complex problems by reasoning about knowledge, represented primarily as if then rules rather than through conventional procedural code.

According to Ignizio (1991) the expert system is a computer aided problem solving system, which is generally based on several artificial intelligence methods, contains the knowledge of a special problem area (domain), helps in analyzing and solving large, complex problems considering the problem solving process of a human expert. So expert systems extend possibilities of computers and make easier computer programming. Expert systems reveal the given problem and suggest a possible solution by coding of the knowledge and the inference possibility of a human expert.

Knowledge based System (KBS) can be effective purely for the purpose of guided concept discovery. KBS is a computer program that reasons and uses a knowledge base to solve complex problems. The term is broad and is used to refer to many different kinds of systems. The one common theme that unites all knowledge based systems is an attempt to represent knowledge explicitly via tools and rules rather than implicitly via code the way a conventional computer program does.

The difference between Knowledge based system and expert system is, knowledge based systems are mainly interested in categories, and expert systems are interested mainly in decisions. Knowledge based expert system corporate these two difference of knowledge based system and expert system to reduce problem of road traffic accident.

KBES is an intelligent computer program that uses the knowledge and inference procedures of human experts to solve difficult problems, and it is not robustly studied in the field of transportation. Some KBES projects are achieved by the several states Department of Transportation (DOT). Their concerns are mainly in Pavement Management System and some others. (Harris et al …, 2005)

Knowledge Based Expert System (KBES) is a collection of Artificial Intelligence (AI) techniques that enable a computer to assist people in analyzing specialized problems. KBES provides human expertise through both the knowledge engineering language and the program supporting environment (King et al.., 2002).

The aim of this study is to develop KBES for decision making on the complex problem of road traffic congestion and accident in case of Jimma town based on the common cause in knowledge base. The KBES developed is a menu driven type using open source programs with three very important web development tools such as Apache, PHP, MySQL, with "C" program language scripting. WEKA data mining technology is also used for traffic accident knowledge acquiring; determine representation of sets of rules for acquired knowledge and modeling of represented knowledge.

Accordingly the system had able to present mitigating measures and determines potential alternatives reduction for accidents. The occurrence of these problems can be investigated with research of countermeasures that were provided by experts on well documented traffic data record references in the past years. Ideas or knowledge of these experts can now be transformed into a computer based knowledge system for end-users' guide and reference.

KBES for Road traffic accident help the user to reach on the right possible solution after accident among road safety maintenance, driver advising, pedestrian recommend/advise, and future care prediction tasks to improve decision on road traffic accident control.

1.2. Statement of Problem

Road traffic accidents (RTA) is one of the major causes of mortality and morbidity around the world and it has been discovered that low and middle income countries are the most affected (Krug et al, 2000). Similar road traffic accident problems are occurred in Ethiopia and all sub cities are affected by these problems.

Road traffic accident management and control has long been a complex issue and seems likely to continue to be so all over the country. Road traffic situation at current time in Jimma is facing difficulty while accident management that is making relevant decision after accident. Because decision makers or road traffic data analysts need to analyze and absorb a large quantity of information about an accident to reach on the right decision at the right time. To some extent the traffic agency tried to reduce road traffic accident using different mechanisms with facing this difficult.

As traffic expert explained and traffic report of Jimma town shown the combination of factors has immensely contributed to the incidence of road traffic accidents (RTAs) in Jimma. These factors includes driver driving speed, detect of vehicles, poor road signals, congestion, poor awareness of pedestrian and poor traffic regulation enforcement. Jimma traffic agency report also shows that from time to time road traffic policemen agreed to reduce the accident using mechanism. But they fail to do so because of increasing accident and unable to manage accident based on the data in suitable way of reminding decision once transferred.

Various studies have addressed the different aspects of road traffic accident controlling mechanism, road safety, mining road traffic data, etc with most focusing on predicting or establishing the critical factors influencing injury severity (Chang, et al.., 2005). Also numerous data mining related studies have been undertaken to analyze road transportation authority data locally and globally. From these studies frequently results are reached depending on the socio economic conditions and infrastructure of a given location. Accordingly to recommendation of Chang (2005) and other local researcher, decision support system is better to control and reduce road traffic accident in Ethiopia. The reason behind this recommendation is the current work is manual; the human brain is capable of evaluating only a very small set of information by incorporating the variables involved in road traffic accident management and control.

This study deal with development of decision making and decision support system from the recommendation of previous researchers. Large number of factors involved in road traffic congestion and accident used as important part of this research. The researcher has been confident in system able to avoid incorrect decision been taken by decision maker while using manual system. For example, avoid situation like; increasing the number of traffic policemen if accidents occurred infrequently and most drivers obeyed traffic laws from signals. So it is important to develop a knowledge based expert system to help decision makers for analysis of traffic laws and policies for the right preventive measurement.

In this regard, the study has explored and founds solutions for the following research questions:

- What are the domain knowledge availables for controlling and managing road traffic accidents problem?

- Which system prototype design is best for a domain expert (traffic officer) to uses a system in user friendly manner?
- Does the study addresses current road traffic accident problem in domain area, explore its reasonable solution and address drawback for further study?
- To what extent RTACKBES will perform decision making process, generate preventive measurement to reduce accident and get user acceptance on specified domain area?

1.3. Objectives

1.3.1. General Objective

The general objective is to develop a knowledge based expert system for improving of decision making process for road traffic accident.

1.3.2. Specific Objectives

- ✓ To acquire knowledge from domain experts and relevant documents for extracting the domain knowledge used in the real life.
- ✓ To design best system user interface that can help traffic officer to use and access system with user friendly tools and techniques.
- ✓ To develop KBES prototype that make decision on road traffic accident problems and have good performance of usage.
- ✓ To evaluate system performance among solving current road traffic accident problem, predict forward research and conclusion.

1.4. Scope and Limitation of the study

The main tasks involved in this research are identifying the knowledge available in road traffic accident management and control, analysis road traffic accident information in term of cases and illustrating decision by system. This research involves end users providing the right decision for the case inputted by traffic expert. This system was included the following assumption:

- ✓ It is assumed that traffic officer is knowledgeable on basic road traffic related problems upon observing congestion and accident location patterns and use of computer application.

- ✓ The study has dealt with road traffic accident concerning vehicle to vehicle crash accident, vehicle to pedestrian crash accident.
- ✓ For this researcher use expert system programming language such as apache, MYSQL as a knowledge base to collect facts, rule and its represented knowledge; and PHP to manipulate rule and execute the code written using C scripting language for the development of road traffic accident control.

The activities not considered in this study are:

- Automobile mechanical error will not be considered as causes for traffic congestion and accidents
- Private sector end users (such as Insurance Company) could not use this system to address traffic accident for their purpose relate to accident. Because generated solution by the system are measured as how much the injury is severe, is it important to make preventive measurement and what final solution will be directed by traffic officer. Private sector but perform all operation related to payment for their insurance member and injured body.

1.4.1. Limitation

The main constraint that the researcher faced while doing this research was data quality and size. Jimma traffic office includes data with the hard copy format in afan oromo language. The researcher taken a difficult time to convert all data from hard copy to dataset and from afan oromo to English language. From 2013 to January 2017 more than 5000 accident data are recorded with 31 attributes on paper. From these instance 1863of recorded instance are unique and most them are similar accident data. The researcher face challenge in identifying the proper instance for selected 19 attributes on dataset.

1.5. Significance of the study

Developed prototype for road traffic accident management and control uses rule and logical techniques to provide alternative potential preventive measurement for traffic decision makers. Tools of Logic allow traffic police officer to measure imprecise and dynamic factors and to arrive at a reasonable judgment like preventive measurement. This system allows end user and

stakeholders making the right decision at the right time with precise and correct output report on road traffic congestion and accident problem.

In addition Knowledge Based Expert System will serve as a tool for road traffic accident management and control in improving conditions at road traffic congestion and accident. The system also encourages users in the adoption of the traffic accident management control alternative. Finally, the importance of this study for Jimma town as well as other city up to the country level helps all traffic officers to simply evaluate road traffic accident inputting the cases or factors occurred during accident, deliver on the right decision and decide alternative preventive.

1.6. Research Methodology

To achieve the main objective of this study, to determine the research design and reach on effective result different methods and techniques are used.

1.6.1. Literature Review

Relevant literatures such as books, journal articles, conference proceedings, manuals, and resources from the internet are reviewed for achieving the research objective. This help the reader of this research to clearly understand the general idea and problem related to road traffic accident. It also illustrated proper method, approach and schema to collect necessary data related to knowledge based expert system due to complete this study successfully.

1.6.2. Study area

The study area of this study is Jimma town. Researcher investigated this research on Jimma town due to optimization of road traffic accident in Jimma. Increasing number of vehicle (motors) and pedestrian optimize injuries severity of accident. There were different research related to road traffic accident was conducted specially performing data mining result to mine accident data, predict injury severity, etc in Ethiopia (Bash et al…, 2012). Thus authors recommended decision support system for future researchers because data mining application alone have no power in problem handling process and contribution. With replying this recommendation the researcher of this study decided to develop RTACKBES on Jimma area. There are two alternative traffic

police offices or stations independently in Jimma and the total of them are four (4). Total numbers of domain are two per station and other traffic police experienced in solving traffic accident problems were included. The total number of traffic police found in Jimma town from both station are about 49 but the researcher interview with one traffic expert from each station purposely to gain road traffic accident decision making knowledge and experience.

Domain experts were selected by Purposive sampling techniques. Burns and Grove (2001) have defined purposive sampling as "Judgmental sampling that makes the conscious selection of certain subjects or elements to be included in the study". Domain experts are selected based on their educational qualifications related to the domain area, year of experience and willingness (Tongco, 2007; Palys, 2008).

In this study, the researcher were taken domain experts that are working as reporter and officer from Jimma town traffic police purposively for interview, because the purposive sampling technique is also used to acquire the required domain knowledge from the limited number of available domain experts.

1.6.3. Research Design

The method a researcher follows to design road traffic accident management and control research is empirical research design. According to Goodwin (2005) Empirical research is research using empirical evidence. It is a way of gaining knowledge by means of direct and indirect experience. As a result, in this study the researcher used experimental method for model building, analysis, and prototype development and testing, whereas non-experimental method was used for knowledge elicitation through discussion with experts and document review. It is a collection of research designs which use road traffic accident knowledge manipulation and controlled testing to understand causal processes. More road traffic variables are manipulated to determine their effect on a dependent variable (decision making) and preventive measurement process.

A researcher use Knowledge Engineering approach or method for KBES development. It is based on the debate on symbolic and situated models of human cognition and aims to make knowledge engineering a learning process by adherence to incremental development,

participation of the expert and direct application of the knowledge based system (Schilstra and Spronck, 2001). In this study the researcher will select Knowledge Engineering approach. Because the Knowledge Engineering approach help the researcher to process road traffic congestion and accident data using steps that are acquire, transfer and represent the experts' knowledge in form of computer system. Also according to Sarma et al. (2009), the main procedures or steps such as identification of input problems, knowledge acquisition, conceptual knowledge modeling, knowledge representation and testing and evaluation are followed in Knowledge Engineering approach.

1.6.3.1. Knowledge Acquisition

The researcher used both primary and secondary source of data. The primary data include the data gathered through interview and observation. For this study interview is particular primary technique used. Whereas secondary data includes traffic accident records, train manuals of traffic control and a Journal article were used. The traffic accident data is particularly selected one.

Semi-structured interview techniques were employed to acquire the required knowledge from the selected domain expert. It allows the interviewer to change the order of the questions and add new question based on the participant response. Therefore, this interview focuses on the concept, procedures, guidelines and experience which domain expert used while diagnosing road traffic accident happened.

The researcher used knowledge acquisition process incorporates typical fact finding methods like interviews, record reviews and observation to acquire factual and explicit knowledge. For this study it is very necessary to acquire tacit and explicit knowledge which is important for the development of road traffic congestion and accident system prototype. Even though, there is no powerful method for knowledge acquisition, interview and observation are the most popular methods. Therefore, both primary and secondary data was needed to acquire the required domain knowledge in this study. Since four traffic experts have the same level of knowledge, experience and skill primary data (tacit knowledge) was acquired from two traffic experts, one expert from station1 and another expert from station2 of Jimma traffic police offices by using interview. Critique knowledge elicitation methods of data mining tools are used to filter the acquired

knowledge. The researcher used both structured and semi structured types of interview to discover relevant tacit knowledge and for further consultations throughout the study. Similarly, secondary sources of knowledge such as Jimma traffic manual report or documents, training manuals on traffic controlling system and journal articles were acquired by using document analysis technique.

The researcher used Knowledge Discovery in Database (KDD) process model to automatically acquire knowledge from the road traffic accident dataset using Waikato Environment for Knowledge Analysis (WEKA) version 3.7.5 data mining tool. KDD is an interactive and iterative process, comprising a number of phases requiring the user to make several decisions. Generally, there are five steps in the KDD process (Two Crows Corporation, 1999; Azevedo & Santos 2008).

Data selection: This stage consists on creating a target dataset, or focusing on a subset of variables or data samples, on which discovery is to be performed. The data relevant to the analysis is decided on and retrieved from traffic accident data collection on the paper. The researcher selects the data set from either traffic police office or station in hard copy format. The collected data with hardcopy format was converted into a softcopy in micro soft office excels spreadsheet. The dataset was formalized from spreadsheet by using comma delimited line CSV. The researcher used recorded accident data from January 2013- February 2017 which are prepared for the purpose of reporting the accident with its possible causes, solution and measurement. The record collected from traffic police station1 was 927, and station2 was 706. The total record collected was 1863. These 1863 unique data are selected from more than 5000 traffic data available in traffic office of mostly similar accident data.

Data pre-processing: This stage consists on the target data cleaning and pre-processing in order to obtain consistent data. Recorded data on paper was written in to Microsoft office excels. The researcher used dataset from a flat file or a spreadsheet. The researcher performs other preprocessing activities to make the data more suitable for data mining like data cleaning, removing attributes and handling missing values while determine model.

The data selected was resampled with percentage split. As a result from 1863 accident data recorded 543 data was selected for the knowledge base through WEKA data mining techniques to determine rules. Filtering algorithm used for removal of inconsistent instance was unsupervised (remove Frequent Value). For all this instances there were 31 attribute on the paper originally and 19 of attribute were selected for the most similarity of attributes. Traffic expert mention the researcher that 19 attribute are most enough to make decision and 12 attribute are unnecessary attribute that needs removal. Removed attribute was zone, city, woreda, kebele, etc which are represented with particular area an accident happened.

Data transformation: It is also known as data consolidation; in this phase the selected data is transformed into forms appropriate for the mining procedure. This stage consists of the transformation of the data using dimensionality reduction or transformation methods. As a result makes data transformation for some selected attributes to make the data more suitable for data mining within different ranges. The data with hard copy format transformed to the spreadsheet form by the researcher. Then it is transformed to the form of dataset that is suited for the data mining. The data is processed by using WEKA data mining and saved in the form of ARF file format. These data are extracted in the form of knowledge by the rule generated using data mining experiment algorithm Finally extracted knowledge coded in knowledge base by C scripting language on MySQL by PHP .

Data mining: It is the crucial step in which clever techniques are applied to extract potentially useful patterns. It consists on the searching for patterns of interest in a particular representational form, depending on the DM objective. The researcher used classification technique on road traffic accident data set which have been collected from Jimma traffic police to develop a model that can predict the possible solution, determine preventive measurement so that to use the model for knowledge based expert system development. Classification is form of data analysis that can be used to extract models describing important data classes or to predict future data trends and classification predicts categorical (discrete, unordered) labels (Asghar & Iqbal, 2009). The researcher conducted two experiments for two classification algorithms namely J48 pruned, and PART. These two algorithms help researcher to achieve on the total number of attribute (to identify selected and rejected attribute) and instance correctly or incorrectly classified. For

conducting this research the WEKA (Waikato Environment for Knowledge Analysis) version 3.7.5 (for windows OS) DM software is chosen. WEKA is chosen because of its widespread application in different DM researches and familiarity of the researcher with the software. WEKA (Waikato Environment for Knowledge Analysis) is a popular suite of machine learning software written in Java, developed at the University of Waikato, New Zealand. WEKA is free software available under the GNU General Public License. The WEKA workbench contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to this functionality (Khac & Kechadi, 2010). It is written in Java and runs on almost any platform. The algorithms can either be applied directly to a dataset or called from your own Java code. The data provides any meaningful information that can be used to know anything about any object (Ionita, 2011).

Interpretation/Evaluation: This stage consists on the interpretation and evaluation of the mined patterns. 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. In DM evaluation serves two purposes. First, it helps to envisage how well the final model will work in the future (or even whether it should be used at all). Second, as an integral part of many learning methods, it helps to explore the model that best represents the training data. 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. Likewise, classification models that are developed in this research are evaluated using a test dataset based on their classification accuracy and interpretation also made accordingly. As a result a test instance which registers more than 90 % accuracy was taken as a knowledge source for KBES development.

1.6.3.2. Knowledge Representation

After the knowledge is acquired, The researcher represented it by using rule based knowledge representation method. For this research the knowledge representation method, rule based, is chosen because it clearly demonstrates the domain knowledge. In a rule based system, much of the knowledge is represented as rules, that is, as conditional sentences relating statements of facts with one another. As a result, rule based representation method is more appropriate to represent and demonstrate the real domain knowledge in applying road traffic accident control system. These rule based representations are extracted from data mining experiment result of PART rule induction algorithm automatically. After knowledge is acquired from domain expert and accident data recorded on paper, the rule means if… (Premises) then … (conclusion) rule are extracted from PART rule induction algorithm of WEKA and RTACKBES developed using web development tools. The fact rules are collected in knowledge base MySQL, Inference Engine performed knowledge chaining and matching for the final solution by C scripting language and the system is compiled using PHP.

In RTACKBES certain knowledge entities, such as pathways and guidelines, are expressed as rules and nature of the accident will force the researcher to use rule-based knowledge representation method since the diagnosis and the treatment of road traffic accident are full of rules and decision tables which can be easily converted to rules by using induction rule method.

The acquired knowledge from the domain experts was represented by using decision tree modeling in a formal language logic which is suitable for the selected algorithm. Rule based reasoning mechanism were employed for the inference engine. In knowledge based system there are many reasoning mechanisms; among that the most commonly used are rule based approach, case based approach or the combination of the two. Case based approaches are designed to work in the way that the basic idea of similar problems having similar solutions (Aboneh, 2007). It is a rule based system that solves problems by remembering past situations and reusing its solution and lesson learned from it. Rule based reasoning, on the other hand reason from domain knowledge represented in a set of rules. The basic format of a rule is:

IF

<<Condition>>

THEN

<<Conclusion>>

Where represents premises and represents associated action for the given premises (Henok, 2011).

1.6.3.3. Knowledge Modeling

A knowledge based system (KBS) can be viewed as a system composed of interrelated elements (Reynaud et al.., 1987). Experience has shown that eliciting and explicating knowledge is best seen as a modeling activity, also called conceptual modeling. This activity takes the form of a specialized type of requirements engineering. KBES construction methods typically provide tools for knowledge analysis in the form of so called conceptual models of knowledge or simply knowledge models.

In this study knowledge discovery or data mining models such as J45 decision tree (decision table) algorithm for the construction of chained knowledge and PART algorithm to define rule for knowledge representation.

According to Vadera (2005), the knowledge model consists of a set of decision tables. The decision table allows visual representation of a number of related rules. In a decision table, knowledge is organized in a spreadsheet format, using columns and rows. The table is divided into two parts. First, a list of attributes is developed, and for each attribute, all possible values are listed. Then a list of conclusions is developed. Finally, the different configurations of attributes are matched against the conclusion. Knowledge for the table is collected in knowledge acquisition sessions. Once the table is constructed, the knowledge in the table can be used as input to knowledge representation methods.

1.6.4. System development methodology

Prototyping approach is followed to develop the Knowledge based expert system. Prototyping allows participating users and domain experts for evaluating systems performance and efficiency.

1.6.4.1. Implementation Tool

The knowledge based expert system was implemented by using web-based, open-source, and cross-platform. The reason is that web technology is used for the reliable user interface for user needs. It is suited for traffic policemen able to use system efficiently, help to achieve the right decision on accident problem, achieve on potential alternative preventive measurement, prepare report and print outcome in a user friendly manner. Also the factor (environment, vehicle, weather, pedestrian, driver, road, etc) data can be updated as much the increasing road traffic accident formulation. Mainly this web application is chosen for displaying prediction of severity, preventive measurement, determining the probable risk of current accident in future with the alternative control.

The developed system for road traffic accident management and control were used PHP script by using the three most important web development tools: The APACHE to provide database, MySQL to handle real time data, and PHP to process the rule and decision. PHP is a server side scripting language for making logic driven websites. Much of its syntax is borrowed from **C,** Java and Perl with a couple of unique PHP specific features thrown in**.** The web server can be compared to a waiter to which if you request something, it can return something prepared on demand such as ham sandwich etc. the application layer is the cook who accepts request from the waiter to prepare the ham sandwich and may need other sources to get information on how to prepare it. The database server is the cook book that stores recipes and organizes them by categories and indices. Rule which is used in these tools is if… else… that is determined and written by the investigator.

Knowledge base provides organized road traffic data saved in spreadsheet, converted to CSV dataset and Data mining tool WEKA version 3.6.5. WEKA is chosen since it is proven to be powerful for data mining and used by many researchers for mining task and the most of researcher are familiar with the tool. It contains tools for data preprocessing, clustering, regression, classification, association rules and visualization. For selection of proper knowledge representation in MYSQL database classification technique is used.

1.6.5. Evaluation

After developing knowledge based expert system, two evaluation technique researcher used to evaluate performance of RTACKBES was carried out. The first is system performance evaluation by using test cases and the second is user acceptance. In other word it was performance of the road traffic accident control prototype and acceptability by the users. In this study, four traffic police officers from Jimma traffic police station were selected for the purpose of testing the road traffic accident KBES prototype. In addition, another four traffic police have experience in handling traffic accident was test case. So system user acceptance was evaluated by 8 evaluators. The selection of evaluators was based on the knowledge, skill and experience they have in road traffic accident data analysis.

The evaluators could evaluate system comparing the functionality of the developed road traffic KBES developed and the manual system they were using. Accordingly, four (4) road traffic accident test cases circulated equally to the evaluators that are one (1) test case (total of 4 test cases) for each evaluator. In the process of testing the performance of the prototype system, the domain experts was classified correctly and incorrectly diagnosed traffic police request, identify the proper decision, determine traffic accident severity, and make preventive measurement by comparing the judgments that were reached by the prototype system with domain expert's judgments reached on the same traffic police experienced handling accident. The evaluators have been assessed the issues of user acceptance testing by using standards such as simplicity of use and interaction with the prototype system, attractiveness of the prototype system, efficiency in time, the accuracy of the prototype system in reaching a decision to identify preventive measurement of a road traffic accidents, including adequate knowledge in the prototype system, the ability of the prototype system in giving the right decision, the ability of the prototype system to update its rule base at runtime, and the importance of the prototype system in the domain area through visual interaction methods. Different researchers used these evaluation standards (Seblewongel, 2011; Solomon, 2013). Finally questionnaires method was used to gather feedback from evaluators. The questionnaire has included both closed ended and open ended questions.

1.7. Organization of the Thesis

The study is organized into six chapters. Chapter one is the introduction part, which contains the background of the study, problem statement, objectives, scope of the study, the significance of the study and methodology to carry-out the research.

Review of conceptual and related literature on the knowledge based systems, about its background, architecture, development phases, and knowledge based system overview and application areas in computer domain are presented in chapter two.

Chapter three discusses the knowledge acquisition, representation and conceptual modeling procedures

Chapter four discusses about Design and Implementation. In this chapter the design and implementation of the prototype are realized by using Data mining results as a knowledge source. The architecture of the new prototype KBE system for road traffic accident management and control is developed. The implementation tool used is web development tools (Apache, PHP and MYSQL).

Chapter five discusses about Testing and evaluation of the proposed systems. In this chapter the performance of the prototype is evaluated both the performance of the system and the acceptance of the system by the users. In addition, discussion was made to show the significance of the proposed approach with previous researches.

Finally, chapter six focuses on the conclusion and recommendation based on the results of the research finding for further research work in the domain area.

Chapter Two

2. Literature Review

The aspiration for computer systems being able to support human experts during complex problem-solving task is a usual topic of AI research. In order to enable a computer system to give rational support when solving problems in a complex application domain, it is essential to provide it with specific knowledge within that domain. A number of methodologies to realize such knowledge knowledge-based systems have been developed, such as, rule-based approach.

In the mid-1970s one of the most significant accomplishments in the field of artificial intelligence (AI) has been the development of the Knowledge Based Expert Systems (KBES). These systems are interactive computer programs that employ a collection of judgment, experience, rules of thumb, intuition, and other expertise in a particular field, coupled with inferential methods of applying this knowledge, to provide expert advice on the variety of tasks (Gasching et al…, 2014). A KBES provides human expertise through both the knowledge engineering language and the program-supporting environment (Harmon et al…, 2005). The AI/KBES application requires development of a generalized knowledge base that permits traffic policemen to interact with the following components: the traffic characteristic data, the theoretical results, and the specific hypothesis for measuring the effects of traffic control measures.

2.1. Artificial Intelligence:

The term Artificial Intelligence (AI) refers to the activity of building intelligent systems. It is a technology of making computers to simulate human beings intelligence (Raza, 2009). An intelligent system is a system that exhibits and possess some basic attributes such as performing some actions, reasoning about a particular domain, making decision and goal oriented problem-solving capability. A system or an agent can be said to be intelligent when the agent's performance cannot be distinguished from that of a human performing the same task (Honavar, 2006). The prime goal of Artificial Intelligence research is to increase human beings understanding in all aspects like in human being's perceptual,

reasoning, learning, and creative processes (Honavar, 2006). The first major and successful Artificial Intelligence research application technologies are expert systems or knowledge based systems (Pomykalski et al..., 1999). Early efforts in building Artificial Intelligence programs were intended to create general-purpose problem solvers.

Over the years, there have been various application areas that have been successful for expert system development. Durkin (1995) and Liebowitz (1991; 1994a; 1994b) present representative problem areas where expert systems have been successfully built: interpretation; prediction; diagnosis.

2.2. Knowledge based expert system

KBES is part of artificial intelligence which helps to develop prototype for the problem needs human expertise. KBESs is the field of artificial intelligence (AI) with the aim of emulating and imitating social problems and then providing human solving behavior in complex real world tasks. The potentiality of such systems and their possibilities of application to transport problems have generated considerable within the transport engineering field (Yeh et al…, 2001).

A KBES provides human expertise through both the knowledge engineering language and the program supporting environment (Harmon et al…, 2005). The AI/KBES application requires development of a generalized knowledge base that permits user to interact with the problem characteristic data, the theoretical or simulation results, and the specific hypothesis for measuring the effects of problem handling measures.

According to Wentworth (1990) KBESs are described by manipulating symbolic, non numeric information which is a logical representation of the problem domain, and these are based on a clear and explicit separation between the knowledge they store (called knowledge representation) and the mechanisms for manipulating such knowledge (called inference and reasoning strategies). Furthermore, these KBES are provided with some limited form of self-knowledge; that is, they are able to keep a representation of their internal structure and function. This characteristic allows them to control their reasoning and to reconstruct inference paths for the

purpose of explaining and justifying solutions to the user. This point is a set of distinguishing features which characterizes their kind of system. (Waterman 1996 and Wentworth 2000)

In order to understand detail of KBES, it is necessary to define the basic terms: Artificial Intelligence, Expert System, Knowledge Engineering, Representation and Knowledge Based System (Waterman 1996).

2.2.1. Expert system

An expert system (ES) is a software system that incorporates concepts derived from experts in a field and uses their knowledge to provide problem analysis to users of the software. The expert system utilizes what appears to be reasoning capabilities to reach conclusions. There are five major components of artificial intelligence (AI) applications in ES designs: (a) expert system, (b) domain expert, (c) knowledge engineer, (d) expert systems building tool, and (e) end-user.

2.2.1.1. Domain Expert:

Domain expert is a knowledgeable person with a reputation for producing good solutions to problems in a particular field.

2.2.1.2. Knowledge Engineer:

The knowledge engineer is usually a person with a background in computer science and AI technology with the purpose on how to build expert systems. The knowledge engineer interviews the domain experts, organizes the knowledge, decides how it should be represented in the expert system, and may assist in the development of a specific program.

2.2.1.3. Expert System Building Tool:

The expert systems building tool is the computer-programming environment and language used by the knowledge engineer or computer programmer to build the expert system.

2.2.1.4. End User:

The user or the end user is the person for whom the expert system is developed and uses the system for its problem solving assistance.

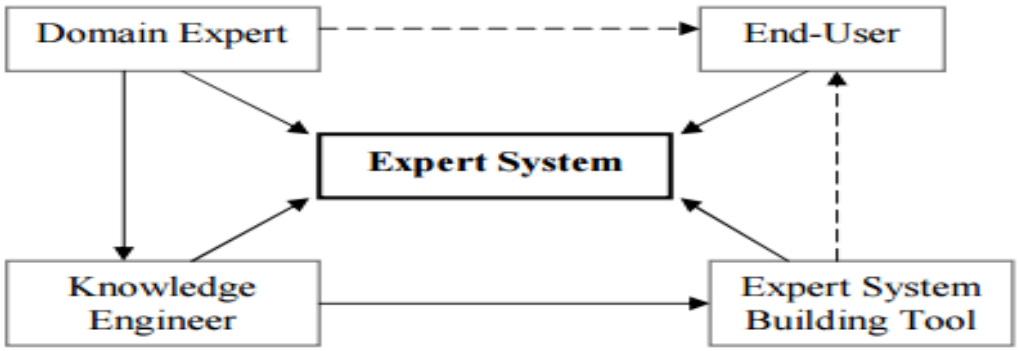

Figure 2.1 Main components of expert systems design (Pomykalski et al., 1999)

The domain expert may add new knowledge to the existing system while the knowledge engineer refines the knowledge in the system. The knowledge engineer converts a domain expert's specialized knowledge into sets of IF-AND-THEN-ELSE rules using instructions that a computer understands. The domain expert may be the knowledge engineer and this may be of an advantage in the accuracy of the expert system. The user may not be a traffic engineer but it is recommended that the user be a traffic engineer or a civil engineering student who finished a course in traffic engineering. The user seeks advice on the system for decision-making processes that may solve a certain problem.

2.2.2. Knowledge based system

Knowledge-Based System (KBS) is one of the major family members of the AI group. With availability of advanced computing facilities and other resources, attention is now turning to more and more demanding tasks, which might require intelligence. The society and industry are becoming knowledge oriented and rely on different experts' decision-making ability. KBS can act as an expert on demand without wasting time, anytime and anywhere. One may consider the KBS as productive tool, having knowledge of more than one expert for long period of time. In fact, a KBS is a computer based system, which uses and generates knowledge from data,

information and knowledge (Sajja & Akerkar, 2010). These systems are capable of understanding the information under process and can take decision based on the residing information/knowledge in the system whereas the traditional computer systems do not know or understand the data/information they process. The KBS consists of a Knowledge Base and a search program called Inference Engine (IE). The IE is a software program, which infers the knowledge available in the knowledge base. The knowledge base can be used as a repository of knowledge in various forms. This may includes an empty Workspace to store temporary results and information/knowledge pieces/chunks.

According to the classification by Tuthhill & Levy (1991), there are main 5 types of the KBS exist: Expert Systems, Hypertext Manipulation Systems, CASE Based Systems, database in conjunction with an Intelligent User Interface and Intelligent Tutoring Systems.

Expert systems, in which the knowledge of a recognized human expert is modeled, are a subset of KBSs. They enable the formalization, preservation and dissemination of specialist knowledge, which has typically been accumulated over a long period of time. The modeled knowledge is available to be managed and applied as an organizational asset.

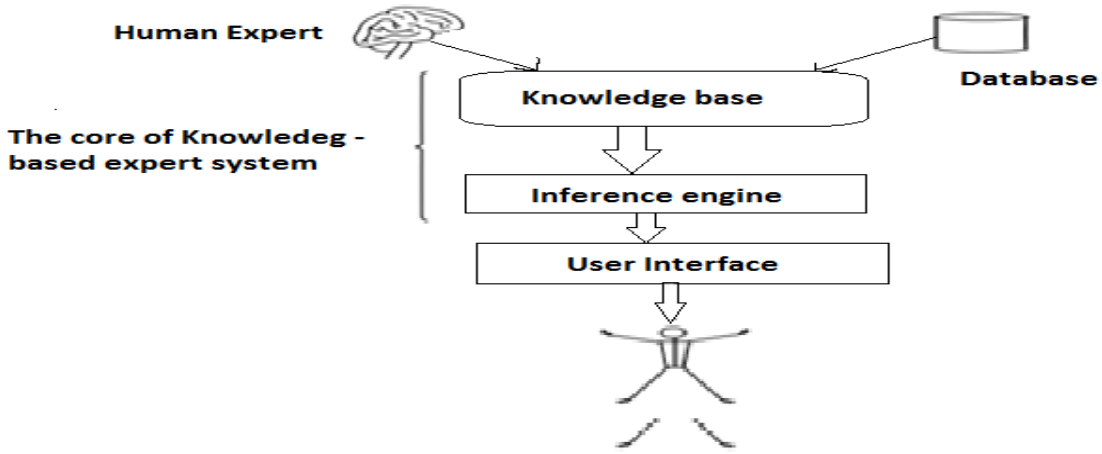

Figure 2.2 knowledge based systems (Tuthhill et al,.. 1991)

2.2.3. Knowledge Based Expert System Tool

The development of a KBES presently involves the cooperative effort between one or more experts who possess the domain dependent knowledge and a knowledge engineer (KE). The KE

elicits the knowledge and uses either an expert system building tool or a general-purpose language to represent and manipulate it. The representation of knowledge in a KBES is dependent on the selection of the tool or language to be used. The KE must make a choice among several available tools before embarking into a major developmental task; the ease of building an expert system depends in part on the choice of the tool.

According to Lenat et al (2003) there are number of languages and tools are currently available for building KBES. These tools can be grouped into three categories. These are: General Purpose Programming Languages, General Purpose Representation Languages and Domain Independent Expert System Frameworks.

Table 2.1 KBES tools and language (Hayes-Roth, et al…, 2003)

Language and tools	**Description**
General Purpose Programming Languages,	✓ Implemented in a high-level language. ✓ need powerful abstraction mechanisms with which other higher level constructs can be built so as to make programming flexible and easy ✓ Need some novel features, such as facilities for experimentation with large chunks of knowledge, tentative modifications, planning and reasoning strategies. ✓ LISP and PROLOG
General Purpose Representation Languages	✓ programming languages developed specifically for knowledge engineering ✓ not restricted to implementing any particular control strategy, but facilitate the implementation of a wide range of problems ✓ SRL, RLL , KEE, OPS5 and OPS5
Domain Independent Expert System Frameworks	✓ provides the system builder with an inference mechanism, from which a number of applications can be built by adding domain specific knowledge ✓ provide knowledge-acquisition and explanation modules to simplify the construction of the expert systems ✓ Control strategies are restricted to those provided in the original system. ✓ EMYCIN, KAS HEARSAY-III, EXPERT and KMS

The expert systems building tool is the computer-programming environment and language used by the knowledge engineer or computer programmer to build the expert system. End User is the user or the end user is the person for whom the expert system is developed and uses the system for its problem solving assistance. The knowledge engineer converts as domain expert's specialized knowledge into sets of IF-AND-THEN-ELSE rules using instructions that a computer understands. The domain expert may add new knowledge to the existing system while the knowledge engineer refines the knowledge in the system. The domain expert may be the knowledge engineer and this may be of an advantage in the accuracy of the expert system.

Apache and PHP: (recursive acronym for PHP: Hypertext Preprocessor)

Apache is a free software/open source web server. Web server is a computer program that is responsible for accepting HTTP requests from clients (user agents such as web browsers), and serving them HTTP responses along with optional data contents, which usually are web pages such as HTML documents and linked objects (images, etc.) HTTP (hypertext transfer

PHP is a widely-used open source general-purpose scripting language that is especially suited for web development and can be embedded into HTML. PHP is a server side scripting language for making logic driven websites. "PHP is an HTML-embedded scripting language. Much of its syntax is borrowed from C, Java and Perl with a couple of unique PHP-specific features thrown in. The goal of the language is to allow web developers to write dynamically generated pages quickly (http://www.cs50.net/resources/, 2008).

2.2.4. Application of knowledge Based Expert Systems

According to Wentworth 1993 KBES in Transportation which deal the feasibility of using expert systems has been proven for highway applications. This is clearly demonstrated by existing operational systems such as FRED (Freeway Real time Expert System Demonstration), a component prototype real time expert system for managing nonrecurring congestion on urban freeways in Southern California, or ERASMUS, an expert system for pavement assessment and rehabilitation which is operational on thirty five sites in France. Other developed systems such as FASTBRID, (fatigue Assessment of Steel bridges), and the WZTS (work zone Safety training system) show that fully integrated decision aid / training systems are both possible and practical.

With the large numbers of the current senior professionals approaching retirement age, expert systems can perform a useful role in the near future. In 1989 the Organization of Economic Co-operation and Development (OECD), in recognition of the possible future role of expert systems in roadway engineering and operations, initiated a project to determine to what extent, and in what areas expert systems can be utilized in road and road transport research, planning, engineering and management. (WENTWORTH 1993). The systems included in the OECD survey are classified by function in four very broad groups as follows:

- ✓ Road traffic accident management and control and Control - systems developed to advise or assist with road traffic accident management and control and control operations, such as diagnostics of traffic problems from sensor data, incident detection, signage, and signalization;
- ✓ Traffic Impact and Safety - systems for evaluating ways of reducing the impacts of traffic: noise control, safety workzone layout, accident investigation, etc;
- ✓ Highway Design and Planning -systems designed to assist with roadway design, and to analyze roadway needs and problems, such as geometries, landslide forecasting, and drainage; and
- ✓ Highway Management- systems to assist with roadway maintenance and operation and decision making, including pavement maintenance, bridge deck repair, and bridge painting strategies.

A number of authors (Eg Wigan; 1993; Logie and Neffendorf 1994) have drawn attention to the possibility of applying expert systems to aid the study of transport problems. The following table 2.2 summarizes potential applications of expert systems considered in the road traffic accident management and control.

Table 2.2 application of KBES in transportation (Bonsal, Kirby and Kwan 1996)

Application Area	**Use**
Design systems	- Infrastructure design - Network/ junction design - Schedule design - Questionnaires design

Diagnostics and prescriptive system	- Road safety systems - Road maintenance - Structures and equipment maintenance
Control system	- Traffic congestion diagnosis - Roadway safety diagnosis - Hazardous material transportation operations - Air traffic control - Road traffic signal timing control - Dispatching and scheduling
Policy Support System	- Multi-criteria decision making - Treatment of uncertainty - Consistency of policy

2.2.5. Basic Structure of Knowledge Based Expert System

A knowledge based expert system is divided into three basic modules; the knowledge base (state memory), inference engine (control strategies or inference mechanisms) and short term (dynamic) memory called context. The other three additional components, which are important in developing usable and widely accepted systems, are the knowledge acquisition module, the explanation module and the user interface. These modules and their interrelationships are shown in Figure 2.3 below (Waterman D.A. 2006 and Wentworth J.A. 2001)

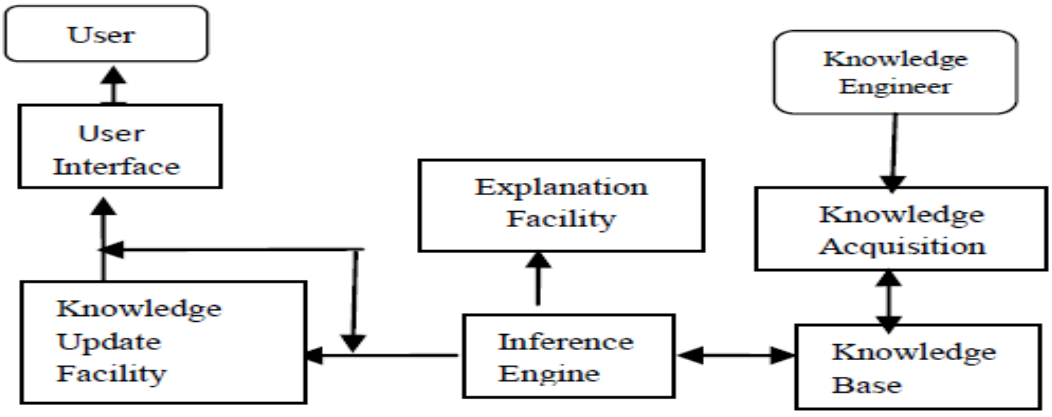

Figure2.3 Basic structure of knowledge based expert system (K P Tripathi 2011)

The expert system building tool can be categorized further as knowledge base, inference engine, working memory, and explanation module as shown in Figure below.

2.2.5.1. Knowledge Base

The knowledge base serves as the storage place for the system's domain specific knowledge needed for understanding, formulating and solving the problem. If contains the facts (data. base) and rules or other knowledge representations (heuristics) that direct the use of knowledge to solve a given problem. So it represents the power of the expert system. Common strategies for representing knowledge that are mostly used in engineering concepts are semantic nets, frames, and rules.

Rules, the most popular type of knowledge representation technique, provide a formal way of representing recommendations, directives, and strategies. Rules are often suitable when the domain knowledge results from solutions developed through years of problem solving experience. These results may come from paper concepts authored by experts or direct interviews from experts. Rules are generally expressed as conditional (if-then) statements.

2.2.5.2. Inference Engine:

Inference Engine is a brain of expert system. It uses the control structure (rule interpreter) and provides methodology for reasoning. It acts as an interpreter which analyzes and processes the rules. Here two approaches are used i.e. forward chaining and backward chaining. Inference engine is the set of procedures for manipulating the information in the knowledge base to reach conclusions. The actual processing of the knowledge is performed by the inference engine. It comprises the control strategy or problem solving mechanism. It combines information supplied by the user with the rules or with the facts in the knowledge base to advise the user (through the user interface), which is how to solve a problem or to reach a goal, including, what conclusions can be reached or what additional information is required. This point is a major contrast between KBES and conventional programs. So it is called by the brain of KBES.

2.2.5.3. Context (short term memory):

The context contains the dynamic or problem specific knowledge, the user response to questions asked by the system, and other temporary information generated by the system. It is also known as the working memory, the work place or short term memory. Memory Contains all the information derived from the inference process. This information describes the problem being solved, the rules that have been "fixed," and the conclusions derived from them.

2.2.5.4. Knowledge Acquisition Module:

Knowledge acquisition is the accumulation, transfer and transformation of problem-solving expertise from experts and/or documented knowledge sources to a computer program for constructing or expanding the knowledge base. It is a subsystem which helps experts to build knowledge bases. For knowledge acquisition, techniques used are protocol analysis, interviews, and observation. The acquisition module is considered to be a subset of knowledge engineering. The main goals in knowledge acquisition module are to help the expert educate system. That is, minimize or eliminate the role of the knowledge engineer as interpreter of the expert's knowledge. Furthermore it is to add in the development and maintenance of the knowledge base.

2.2.5.5. Explanation Module:

The explanation module gives the system the ability to explain its reasoning process and to provide definitions and other information to the user. Also this module helps the domain expert check the system's reasoning in the system debugging. The Explanation Module contains explanations for every inference made or piece of advice given.

2.2.5.6. User Interface:

The user accesses the system through a user interface which should be friendly so that man and machine can communicate directly and efficiently. Currently actual language processors do not exist, so most user interfaces are implemented as a program-oriented subset of English.

2.2.6. Knowledge Based Expert System Building

The major task of building an expert system is to transfer the expertise and knowledge acquired from one or more expert to a computer program. This task of expert systems developers (knowledge engineers) is to carry out such a transformation and to system can reach the desired level. This step will vary depending on the characteristics of the program, the objectives selected. However, the following stages are normally used in the development of an expert system. (Hayes-roth 2006)

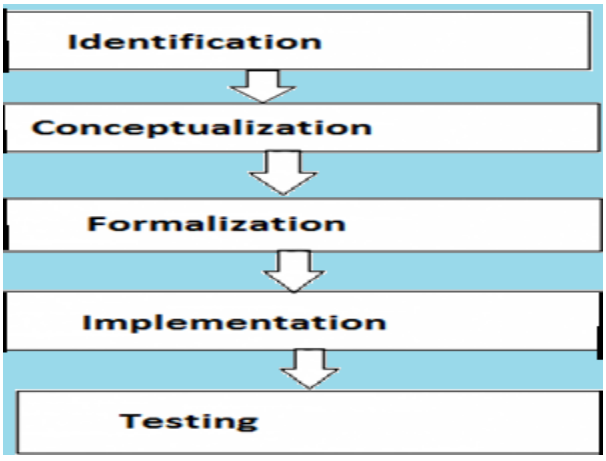

Figure 2.4 Knowledge Based Expert System building stages (Hayes-Roth 2006)

1) Identification

The first step in building a knowledge-based expert system is to identify the area, concepts, and characteristics of the problem and solution. In addition, the participants and resources (time, labor, and computing facilities) needed during development should be identified.

2) Conceptualization

The concepts needed to represent knowledge and the overall structure of knowledge-control strategies must be determined before a preliminary system design can be completed.

3) Formalization

This stage involves design of the formal organization of the knowledge consistent with the development tools or language used. The detailed design of the system is formulated by the experts or knowledge engineers.

4) Implementation

In this stage the knowledge engineer turns the formalized knowledge into a working computer program.

5) Testing

The performance and behavior of the prototype expert system are iteratively evaluated through comparison with the human expert's abilities. Revisions of the system are then made by a knowledge engineer using additional advice from the human experts. In building expert systems, symbol-manipulation programming languages such as LISP and PROLOG have been widely used. These languages have been specially designed for artificial intelligence applications. However, knowledge engineering tool-kit or shell have been also widely used, which comprises and includes an inference, empty knowledge base and context structure, and support facilities such as a knowledge base editor and user explanation facility. Knowledge Based Expert Systems Development: Knowledge engineering

Mostly knowledge engineering, the process of building an expert system, involves some basic steps. The main phases of a knowledge based system development processes are planning, knowledge acquisition, knowledge representation and evaluation (Raza, 2009; Sajja& Shah, 2010).

The knowledge of the expert(s) is stored in his mind in a very abstract way. Also every expert may not be familiar with knowledge-based systems terminology and the way to develop an intelligent system. The Knowledge Engineer (KE) is responsible person to acquire, transfer and represent the experts' knowledge in form of computer system. People, Experts, Teachers, Students and Testers are the main users' groups of knowledge based systems. (Sajja & Akerkar, 2010)

2.2.6.1. Knowledge Acquisition

Knowledge acquisition is the process of acquiring relevant knowledge from human experts, books, documents, sensors, or computer files. The knowledge can be specific to the problem domain or to the problem-solving procedures, it can be general knowledge (e.g., knowledge about business) or it can be meta-knowledge (knowledge about knowledge). Knowledge acquisition is the bottleneck in knowledge based system development today. Because, the

trustworthiness and the performance of the knowledge based system mainly depends upon the acquired knowledge (Alechina, 2012). The knowledge acquisition process incorporates different methods such as interviews, questionnaires, record reviews and observation to acquire factual and explicit knowledge (Gil, 1999). The performance of the expert systems depends upon the reliability, validity and accuracy of the elicited knowledge. The process of knowledge elicitation is affected by different contributing factors such as communication between the expert and ability of knowledge engineer (Tehrani. 2009). Therefore, effective acquisition techniques facilitate to acquire relevant knowledge from domain experts. The commonly used knowledge acquisition techniques are discussed as follows (Wang, 2011).

A. Interview

An interview technique is the process of interacting with domain expert on how they perform their tasks based on their expertise. Knowledge acquired through direct elicitation methods are procedural knowledge. Based on its structure, interview can be classified into structure, semi-structured and unstructured interview.

- **Structured Interviews-** A structured interview method is questioning the domain expert directly. It is goal-oriented process. It forces organized communication between the knowledge engineer and the domain expert. The structure reduces the interpretation problems inherent in unstructured interviews and allows the knowledge engineer to prevent the bias caused by the subjectivity of the domain expert (Ranjan 2006).
- **A semi-structured interview** is an interview which has a guide that usually includes both closed-ended and open-ended questions. It is more flexible than structured one. In these kinds of interview the interviewer has a chance to change the order of questions and expand the dimension of questions based on the participant's responses (Ranjan 2006).
- **Unstructured Interviews-** sessions are conducted informally, usually as a starting point. Unstructured interview techniques provide complete or well-organized descriptions of cognitive processes. There are many reasons that enforced to applying unstructured interview. Domain the experts usually find it very difficult to express some of the most important elements of their knowledge. Through structured interview it is difficult to acquire the required knowledge. With good training and personal experience knowledge engineers

can use unstructured interview to acquire relevant knowledge from domain expert (Ranjan 2006).

Therefore, efficient and effective interview techniques largely depend on the ability of knowledge engineer to articulate their implicit knowledge. Because every interview is different in very specific ways and it is difficult to provide comprehensive guidelines for the entire interview process. Therefore, interpersonal communication and analytic skills of knowledge engineer is very important. On the other hand eliciting knowledge using indirect methods requires human intervention such as observation, document analysis, etc. (Wang 2011).

B. Document Analysis

The final form of knowledge acquisition method is concerned with a detailed analysis of the existing document. This technique is used to collect relevant knowledge from the existed documents of different format. These documents include professional literature, brochures, manuals, guidelines, employee handbooks, reports, glossaries, course texts, and other relevant materials (Gau.1990). Knowledge elicitation methods can be classified into different types. Direct and indirect is the commonly known methods of knowledge elicitation. The way of classification depends upon how knowledge engineer directly obtains information from the domain expert (Osuagwu 2006)]. A direct method involves directly questioning a domain expert on how they do their job. In order to implement direct methods successfully, the domain expert has to reasonably articulate and willing to share his/her knowledge. However, in case of indirect methods the required knowledge is not requested directly. Instead, the result of the knowledge elicitation session must be analyzed in order to extract the required knowledge. Indirect methods are thought to be more suitable when knowledge is not easily expressed by the domain expert (Osuagwu 2006).

Knowledge representation- Acquired knowledge is structured so that it was ready for use in the process of knowledge representation. This activity involves preparation of a knowledge map and encoding of knowledge in the knowledge base.

Knowledge validation- Knowledge validation (or verification) involves validating and verifying the content of knowledge (e.g., by using test cases and confusion matrix) and user acceptance. The testing result of knowledge based system was validated by domain expert.

Inference- This activity involves the design of software to enable the computer to make inferences based on the stored knowledge for the specific domain problem. In other word inference engine is a programs that reason over extensive knowledge bases.

Explanation- This step involves the design and programming of an explanation facility. Explanation module is program that answered how the knowledge based system arrived at certain conclusion. Explanation module addresses the issues of system user interactivity

2.2.6.2. Knowledge Modeling

. Central to the current perception is the knowledge model, which views knowledge acquisition as the construction of a model of problem-solving behavior that is, a model in terms of knowledge instead of representations. The concept of knowledge-level modeling has matured considerably. The practical knowledge-level models incorporated in today's methodologies do not simply reflect the knowledge content of a system; they also make explicit the structures within which the knowledge operates in solving various classes of problems. This enables the reuse of models across applications. The model structures provide a framework for knowledge acquisition and a decomposition of the overall acquisition task. Identified parts of knowledge-level models domain models or problem-solving methods can serve in different systems or in different roles in the same system. The main advantage remains: The knowledge level focuses attention on the knowledge that makes systems work rather than on the symbol-level, computational-design decisions that provide the operational framework. The knowledge modeling methods can be classified into three categories: manual, semiautomatic, and automatic.

Manual methods are basically structured around an interview of some kind. The knowledge engineer elicits knowledge from the expert or other sources and then codes it in the knowledge base. The three major manual methods are interviewing (i.e., structured, semi structured, unstructured), tracking the reasoning process, and observing.

Semiautomatic methods are divided into two categories: those intended to support the experts by allowing them to build knowledge bases with little or no help from knowledge engineers and those intended to help knowledge engineers by allowing them to execute the necessary tasks in a more efficient or effective manner (sometimes with only minimal participation by an expert).

In automatic methods, the roles of both the expert and the knowledge engineer are minimized or even eliminated. For example, the induction method, which generates rules from a set of known cases, can be applied to build a knowledge base. The roles of the expert and knowledge engineers are minimal. The term automatic may be misleading, but it indicates that, compared with other methods, the contributions from a knowledge engineer and an expert is relatively small.

2.2.6.3. Knowledge Representation

To build the knowledge base we have the problem of how to represent it. Knowledge representation concerns the mismatch between human and computer 'memory' .we calls these representations knowledge bases, and the operations on these knowledge bases, inference engine.

A knowledge representation (KR) is an idea to enable an individual to determine consequences by thinking rather than acting, i.e., by reasoning about the world rather than taking action in it .The knowledge acquired from experts or induced from a set of data must be represented in a format that is both understandable by humans and executable on computers. Good Knowledge Representation Languages should be Expressive, Concise, Unambiguous, and Independent of context, Efficient and effective.(Kesarwani & Misra, 2013) Knowledge Representation methods all have advantages and limitations. Production rules are popular in the design of first-generation expert system.

2.2.6.3.1. Frames based Representation

A frame is a node with additional structure that facilitates differentiating relationships between objects and properties of objects. Sometimes it is called as "slot-and-filler" representation. Frames overcome the limitation of semantic network that differentiates relationships and properties of objects. Each frame represents a class (set) or an instance (an element of a class). Frames are application of object-oriented programming for expert systems. The concept of a

frame is defined by collection of slots**.** Each slot describes a particular attribute or operation of the frame. Slots are used to store values. A slot may contain a default value or a pointer to another frame, a set of rules or procedure by which the slot value is obtained.(Sharma & Kelkar)

2.2.6.3.2. Semantic Networks

Semantic networks are an alternative to predicate logic as a form of knowledge representation. The knowledge can be store in the form of graph, with nodes representing objects in the world, and arcs representing relationships between those objects. Semantic network also called as Associative Network. Semantic representation consists of 4 parts:

i. **Lexical**: It tells which symbols are allowed in the representation's vocabulary. Nodes denote objects, links denote relation between objects, and link-labels denote particular relations.
ii. **Structural:** It describes constraints on how the symbols can be arranged. Nodes are connected to each other by links.
iii. **Procedural:** It specifies the access procedures (to create, modify, answer questions). Procedures are constructor procedure, reader procedure, writer procedure and erasure procedure.
iv. **Semantic:** It establishes the way of associating the meaning. Nodes and links denote application specific entities.

Inheritance is one of the main kinds of reasoning done in semantic nets. The ISA (is a) relation is often used to link a class and its super class.

2.2.6.3.3. Case Based Representation

Case-Based Representation is a computer technique, which combines the knowledge-based support philosophy with a simulation of human reasoning when past experience is used, i.e. mentally searching for similar situations happened in the past and reusing the experience gained in those situations (Leake 1996). The concept of case based reasoning is founded on the idea of using explicit, documented experiences to solve new problems. The decision maker uses previous, explicit experiences, called cases, to help him solve a present problem. He retrieves the appropriate cases from a larger set of cases. The similarities between a present problem and the retrieved case are the basis for the latter's selection (Gonzalez and Dankel, 1993).

2.2.6.3.4. Rule Based Representation

Rule based reasoning is a system whose knowledge representation in a set of rules and facts. Symbolic rules are one of the most popular knowledge representation and reasoning methods. This popularity is mainly due their naturalness, which facilitates comprehension of the represented knowledge. The basic forms of a rule, if<condition> then <conclusion> where <condition> represents premises, and <conclusion> represents associated action for the premises. The condition of rules are connected between each other with logical connectives such as, AND, OR, NOT, etc., thus forming a logical function. When sufficient conditions of a rule are satisfied, then the conclusion is derived and the rule is said to be fired. Rules based reasoning was dominantly applied to represent general knowledge. Rule based expert systems have a significant role in many different domain areas such as medical diagnosis, electronic troubleshooting and data interpretations. A typical rule based system consists of a list of rules, a cluster of facts and an interpreter (Prentzas, 2007).

There are two main inference methods in rule based reasoning mechanism. These are backward chaining and forward chaining. The former is guided by the goals (conclusions), whereas the latter one is guided by the given facts (Prentzas, 2007).

During forward chaining, the inference engines first predetermine the criterion and the next steps are to add the criterion one at a time, until the entire chain has been trained. With data driven Control, facts in the system are represented in a working memory which is continually updated.

Rules in the system represent possible actions to take when specified conditions hold items in the working memory. The conditions are usually patterns that must match with the items in the Working memory. In forward chaining, actions are usually involves adding or deleting items to and from the working memory. Interpreter of the inference engine controls the application of the rules, given the working memory. The system first checks to find all the rules whose condition holds true (Shaffer, 1991 & Freeman-Hargis, 2012). Both data driven and goal driven chaining method follows the same procedures. However, the difference lies on the inference process. The system keeps track of the current state of problem solution and looks for rules. This cycle were repeated until no rules fire or the specified goal state is satisfied (Merritt, 2000)

Backward chaining is the strategy focuses its effort by only considering rules that are applicable to the particular goal. It is similar with forward chaining the difference is it receives the problem description as a set of conclusions instead of conditions and tries to find the premises that cause the conclusion. Given a goal state and then the system try to prove if the goal matches with the initial facts. When a match is found goal is succeeded. But, if it doesn't then the inference engine start to check the next rules whose conclusions (previously referred to as actions) match with the given fact. Note that a backward chaining system does not need to update a working memory instead it keeps track of what goal is needed to prove its main hypothesis. Goal driven control is commonly known as top-down or backward chaining

For this result the researcher used forward chaining to add road traffic accident cases in to database with its possible solution.

2.3. Road traffic accident management and control

Road traffic Accidents (RTA) is one of the major causes of mortality and morbidity around the world and it has been discovered that low and middle income countries are the most affected (Krug et al, 2000). The World Health Organization estimates that more than 3,000 people are killed every day in road traffic accidents globally, with at least 30,000 others injured or disabled. This adds up to over 1 million people killed and between 20 – 50 million injured or crippled in road traffic accidents each year (Krug et al, 2000).

Road traffic accident in Ethiopia in its different states like Oromia, Jimma and others are described in EEA, Report on the Ethiopian economy, 2012. The majority of road accident victims (injuries and fatalities) in developing countries are not motorized vehicle occupants, but pedestrians, motorcyclists, bicyclists and Non-Motorized Vehicles (NMV) occupants. The Global Burden of Disease study undertaken by the World Health Organization (WHO), Harvard University and the World Bank showed that in 1990 traffic crashes were assessed to be the world's ninth most important health problem, and by the year 2020 this would move up to third place in the table of leading causes of death and disability.

Traffic accidents have a number of adverse effects on the socioeconomic development of the country. Since most of the victims are those in the productive age group, families lose their bread winners and the country its productive manpower. According to G. Jacobs, A. Aeron-Thomas

and A. Astrop (2000 - 5) hospitals in Ethiopia are stretched with the growing problems of traffic accidents in the country.

Seriousness of road traffic accidents are also reported by concerning body. According to Ethiopian Federal Police (2008/09-2010/11) report, each year more than two thousand people die and ten thousand people being injured because of road traffic accidents. Every year, lots of people die or suffer from accidents because of factors that could have easily been prevented. The growing number of vehicles, ignorance of safety norms and reckless (careless over danger) driving have attributed to the huge number of accidents.

Road traffic control is important issue in all over the world that need attention in order to get gradual solution on the problem among traffic accident. According to Keshav, et al the road traffic accident management and control have long been a complex issue and seem likely to continue to be so. Road traffic laws and policies depend on a large number of factors. By incorporating the variables involved in road traffic accident management and control such as the numbers of accidents, traffic violations, and traffic policemen on duty into an artificial intelligence technique, it is possible to build a traffic decision making system to help decision makers for analysis of traffic laws and policies. (Keshav, et al.., 2005)

Several authors have described decision support systems for road traffic accident management and control. Some of them have used the fuzzy logic technique in their decision process such as. However, all these systems are real time decision support systems, and they do not support the road traffic accident management and control policies (i.e. long term decisions).(**Keshav, et al 2005**)

Table 2.3 Decision support system (Keshav, et al 2005)

Author	**Year**	**Authorize**
S. G. Ritchie	1990	A knowledge-based decision support architecture for advanced road traffic accident management and control
H. Zhang and S. G. Ritchie	1994	Real-Time Decision Support System for Freeway Management and Control
A. Hegyi, B. De Schutter, S. Hoogendoorn, R. Babuska, H. van Zuylen, and H. Schuurman	2001	A fuzzy decision support system for traffic control centers," Intelligent Transportation Systems

K. G. Zografos, K. N. Androutsopoulos, and G. M. Vasilakis	2002	A real-time decision support system for roadway network incident response logistics," Transportation Research Part C: Emerging Technologies,
B. De Schutter, S. P. Hoogendo+orn, H. Schuurman, and S. Stramigioli	2003	A multi-agent case-based traffic control scenario evaluation system," Intelligent Transportation Systems

2.4. Road traffic and knowledge based expert system

Roads traffic may consist of pedestrians, ridden or herded animals, vehicles, streetcars, buses and other conveyances, either singly or together, while using the public way for purposes of travel. Traffic laws are the laws which govern traffic and regulate vehicles, while rules of the road are both the laws and the informal rules that may have developed over time to facilitate the orderly and timely flow of traffic. As a general rule, drivers are expected to avoid a collision with another vehicle and pedestrians, regardless of whether or not the applicable rules of the road allow them to be where they happen to be. (Mann, 2002)

The contribution of human behavior towards traffic accidents is an important area of interest in the remedial attempts to address the global road safety problem. A clear distinction needs to be made between errors and violations to ameliorate traffic accidents, due to the different psychological origins of these components of human behavior, as well as the dissimilar forms of remediation (Campbell et al…, 1990).

KBES is an intelligent computer program that uses the knowledge and inference procedures of human experts to solve difficult problems, and it is not robustly studied in the field of transportation. Some KBES projects are achieved by the several states Department of Transportation (DOT). Their concerns are mainly in Pavement Management System and some others. (Cohn and Harris, 1992)

2.5. Online/ Web based KBSE for road traffic accident control

As a result of kehinde williams (2015) thesis conduct on the Road Traffic Accident Monitoring System (RTAMS) was developed using Adobe Dreamweaver and Notepad++ as the Integrated Development Environments, HTML, CSS and JavaScript were used for the frontend, PHP was used as the scripting language, and MySQL served as the database server. Most of the languages

and tools used were open source which ensured that the application would be robust, reusable, cheap and highly scalable (williams 2015).

Expert systems can assist humans in solving problems that require extensive knowledge or huge amounts of time. These systems are also applicable in dealing with problems related to computer science, agriculture, nutrition, medicine, engineering, education, geology, and so forth. One of their main advantages is that they are easy to access through computer technology (Bianchini, 2012)

Useful KBS and expert systems have already been developed in the field of transportation and safety using web application such as PHP, XML, JS, and Apache. Some KBES researches are conducted by integration of web and windows tools such as JAVA, VB. application using web page as a user interface. Relational database tool such as MySQL are served as a knowledge base for the collection of knowledge and sets of rule. The following list shows some of the examples.

- USLIMITS2 is a web-based expert system which aims to assist engineers in the selection of safe and appropriate speed limits in speed zones on all American roads (Srirnivasan et al…, 2008)
- Paver is a knowledge-based expert system developed for the management, maintenance, and rehabilitation of pavement. Paver is applicable to military installations, municipalities, airports, researchers in universities, and for the use of consultant companies. (Ismail et al.., 2009)
- COPRBU is a knowledge-based expert system developed to deal with problems relating to public buses with respect to routes and schedule, level of service, and their reliability. (Wen, 2008)
- develop web-based advisory expert system for the purpose of applying traffic calming strategies on residential streets is described because there currently lacks a structured framework for the implementation of such strategies. (Falamarzi 2006)

Lack of traffic safety has become a serious issue in residential areas. For this there are different system used in solving this problem. Therefore Web application uses open source technology and necessary to create efficient user interface for the user. It has been proven that web-based expert systems can take on an important role in spreading knowledge among engineers and

researchers because they are accessible anywhere and at any time, only requiring an internet connection with no installation needed. (Lababidi & Baker, 2003)

2.6. Related work

According to Jonas Mwango (2015), in his road traffic accident and possible solution investigation recommended the idea of people on the accident. Despite the early concerns expressed over loss of life, road traffic crashes have continued to this day to exact their toll. The rising trend in number of road traffic accidents is an indicator that a lot needs to be done in terms of instituting effective interventions to reverse the trend. The road safety work has up until now heavily been relying on countermeasures such as regulating and surveillance of behavior, information and education in order to make the road users behave correctly so that accidents will not happen. Most countries are moving away from this initial traditional road-user approach system to a more holistic systemic approach in managing the road traffic system. The traditional approach has been that the individual road-users utterly are responsible when crashes occur and as a result the countermeasures have mostly been aimed at changing the behavior of the road-users in order to adapt them to the road transport system. This intervention has failed as the number of accidents keeps on increasing. Factors leading to serious road trauma attributed to behavior of drivers include speeding, driving under influence of alcohol, driving when tired and the non-use of restraints , other substantial risk factors include the use of mobile telephones, driving under the influence of drugs and failure to observe safe distances.

Various studies have addressed the different aspects of RTAs, with most focusing on predicting or establishing the critical factors influencing injury severity (Chang, et al. 2005). Numerous data mining-related studies have been undertaken to analyze RTA data locally and globally, with results frequently varying depending on the socio economic conditions and infrastructure of a given location. Most of the researchers recommend the future work in case they doesn't achieve on the real solution; developing any decision support system for decision making on the road traffic accident management and control. These researchers investigate their work using data mining on the problems like:

- ✓ Identifying the prediction factors of crashes and crash-related injuries, using models to perform a risk assessment of a given region (Ossenbruggen & Pendharkar et al. 2001),
- ✓ Analyzing accident severity using application of data mining technology (decision tree techniques, knowledge seeker algorism) t in Addis Ababa, Ethiopia (Tibebe 2005),
- ✓ Predicting the degree of drivers' responsibility for car accidents using classification algorithms (Regassa 2009),
- ✓ Demonstrates data mining models for accident severity analysis in support of reducing road traffic accidents by identifying and predicting the major vehicles and driver's determinant risk factors (attributes) that cause road traffic accidents (Mossie 2009,
- ✓ Apply Bayesian Network to express the complicated relationship between the traffic accident and their causes, as well as the correlations between the factors of causes (Hongguo, Huiyong and Fang 2010),
- ✓ Conducted perspective analysis of traffic accident data using data mining techniques. Naive Bayes Bayesian classifier, AdaBoostM1, Meta classifier, PART Rule classifier, J48 Decision Tree classifier and Random Forest Tree classifier were compared for classifying the type of injury severity of various traffic accidents (Hemalatha et al … 2011),
- ✓ Argue significant differences among age and gender groups suggest that drivers perceive and react to pavement surface conditions in very different ways, and this has important safety implications.

In the road traffic accident research conducted above using data mining techniques was not considered as a feature that help user to as a common system to manage and control the road traffic accident happened day to day. It also do not support stakeholders and traffic officer to make decision on the current state of accident. But this study involves the action not considered in the above seven (7) listed works. The researcher concerns the gap and include in his research decision making support feature and system monitor road traffic accident management and control.

Some authors have described decision support systems for road traffic accident management and control. Some of them have used the fuzzy logic technique in their decision process. (Hegyi, et al 2001)

However, all these systems are real time decision support systems, and they do not support the road traffic accident management and control policies (i.e. long term decisions). They realize that making the correct decision at the right time has been a major problem in road traffic accident management and controls, because decision-makers need to analyze and absorb a large quantity of information in a short time. The information can be vague and sometime conflicting in nature. All the above authors investigate data mining application for their study and recommend that knowledge based or any decision support system.

Almejalli et al (2001) investigate a system that can improve road traffic accident management and control policies using Fuzzy Logic. The variables involved in road traffic accident management and control are dynamic and complex: traffic accidents, traffic violations, available traffic policemen, available traffic cameras, etc. There dynamic variables are fuzzified and used to be analyzed with a fuzzy inference system. A fuzzy rules based system is developed using the expert knowledge to generate effective policies. They include if a given city should increase the number of traffic Road Traffic policemen on duty, increase fines, impose stricter speed limits, install better safety devices, or provide better education for drivers. The system itself would not decide the policies, but would provide traffic decision makers with better information so they could make better decisions. In order to measure system success, the proposed decision support system has been evaluated in a number of case studies of Riyadh Region Traffic with a set of real data.

Table 2.4 Selective rule from the Riyadh Traffic Rule-Based structure (Almejalli et al., 2001)

IF	**THEN**
Violations is medium AND Jam is low AND Accidents is very high	Launch Traffic Awareness Campaign is High
Violations is high AND Vs/T-Policeman is high AND Vs/T-Camera is high	Increase Punishment/Fine is Medium
Violations is very high AND Vs/T-Policeman is high AND Vs/T-Camera is very low	Increase Num. of T-Policemen is Very High AND Inc Num. of T-Cameras is Low
Violations is very low AND Vs/T-Policeman is high AND Vs/T-Camera is very high	Increase Num. of T-Policemen is Very Low AND Increase Num. of T-Cameras is Very Low

KBES developers Guzman and Sigua (2009) investigated engineering study guide for evaluating road intersection improvements. According to these investigators recent developments dealt with engineering solutions and eventually decide on how to improve the intersection based on the performance measures results. Programming techniques they used to develop the knowledge based expert system was web-based, open-source, and cross-platform. Their alternative tools were web technology. The program will use a PHP script by using the three most important web development tools: The Apache, MySQL, and PHP. Web applications are commonly used for selling products, publishing information, and web enabling traditional applications. This study deals with the intersection only and doesn't concern about road traffic congestion and accidents.

Another addressed study was by Legien (2015), that he presents an expert System with Web Interface Based on Logic of Plausible Reasoning (LPR). This formalism reflects human ways of knowledge representation and reasoning. The knowledge is modeled using several kinds of formulas representing statements, hierarchies, similarities, dependencies and implications. Several types of inference patterns are defined. The system investigated is called LPR Intelligent Information System (LIIS) which is a web-based application written in Java. It is created with Google Web Toolkit, supporting browser-based application development. The toolkit provides Java API and widgets, which can be compiled to JavaScript frontend. LIIS uses MySQL database for storing knowledge bases (formulas) and system data (e.g. user profiles, privileges).

Therefore the research is necessary in all categories of road traffic congestion and accident. The research gap between this study and the previous investigated in Ethiopia; there is no decision support system such as knowledge base expert system, but recommended by different researchers. Those researchers investigate data mining application for the road traffic accident severity prediction by increasing number of factor up to date. This kind of research is used only to predict the probable of severity less severe, high severe and extremely severe. But not help to make road traffic congestion and accident preventive measurement and could not apply corrective strategies with potential traffic control alternatives.

This study will solve a problem with a system accepting lots road traffic congestion and accident fact from traffic police officer and determine the cause of the accident and predict what kind of measure must taken by driver and traffic

Therefore, in order to improve the gap of previous studies and considering the vital and critical role of a KBES, this study aims to develop Adaptive-learning KBES that can improve the accurate advice that is provided by the system in line with accurate diagnosis and control road traffic accident, and also used to reduce an accident.

The road traffic to be observed and investigated will cover all traffic accident incidents that took place. Traffic control measures will now be analyzed by using rules for inference (drawing conclusion) by forward chaining to achieve potential alternatives.

Tools used for implementation of the expert system are a web-based, open-source, and cross-platform. The following web technology was used. The program will use a php script by using the three most important web development tools: The Apache, MySQL, and PHP. These web applications have one thing in common and that it is composed of a web server, application layer and a database. In this study the mode for KBES development for traffic congestion and accident control is decision tree structures that are the bases for the development of prototype knowledge based expert system. The prototype follows the same procedures as presented in the decision tree when diagnosing road traffic accident problem.

Rule Based Reasoning Techniques an approach which have a numbers of good features expressed as compact representation of general knowledge. Rules can easily represent general knowledge about a problem domain. Therefore, the best alternative technique is rule based.

There are two main inference methods in rule based reasoning mechanism. These are backward chaining and forward chaining. For this research forward chaining strategy was used. Because when you have a small set of initial facts; and when there is lots of different rules which allow you to draw the same conclusion it is better to use forward chaining. Backward chaining may be better if you are trying to prove a single fact, given a large set of initial facts.

CHAPTER THREE

3. KNOWLEDGE ACQUISITION, KNOWLEDGE REPRESENTATION AND CONCEPTUAL MODELING

Knowledge acquisition, representation and conceptual model are a guiding, achieving and implementing technique in development of knowledge based expert system. The researcher follows these techniques to develop knowledge based expert system for analysis of road traffic accident control process and to improve decision making process on road traffic accident.

3.1. Knowledge Acquisition

The first objective of this study is to acquire knowledge from domain experts and data mining of relevant documents for extracting the domain knowledge used in the real life. Knowledge acquisition is the first step to achieve this objective.

Knowledge acquisition (KA) is the process of acquiring relevant knowledge from domain experts and other sources of information such as books, databases, guidelines, manuals, journal articles and computer files. KA is the process of eliciting, structuring and representing (formalizing) domain knowledge acquired from different sources. The term knowledge acquisition and knowledge elicitation have been used interchangeably in the field of artificial intelligence (AI) literature. The acquired knowledge can be specific to the problem domain, it can be general or it is meta-knowledge (knowledge about knowledge). Knowledge acquisition is the first step and time consuming task in the development of knowledge based system (Y. Bassil 2012).

The development of an efficient knowledge-based system (KBS) involves the development of an efficient knowledge base that has to be complete, coherent and non-redundant. The step of knowledge acquisition is one of the major bottlenecks in the stage of case base development. Usually, for each application domain there are several sources of knowledge (human experts, the specialized literature which includes textbooks, books, reviews, collection massive of data, etc) (Asghar & Iqbal, 2009).

According to Asghar & Iqbal (2009), in order to make knowledge extraction as much as correct as possible (i.e. in order to keep the correctness of the knowledge as it is kept at the source) different techniques could be applied. Among these techniques, data mining techniques and, more general, knowledge discovery techniques became the most used in the recent years.

There are certain important steps that the knowledge engineer need to carry out during knowledge acquisition process (Aboneh. 2007):

i. Eliciting data and information from the domain experts
ii. Interpreting the acquired information to understand human expert reasoning processes
iii. Constructing model to structure the expert‘s knowledge
iv. Represent knowledge and build knowledge base
v. Repeating step I-IV as the knowledge base system involves into a functional system

3.2. The Process of Knowledge Acquisition

For the purpose of this research, knowledge engineer acquire knowledge using the process of knowledge acquisition includes some basic activities such as interviewing of domain expert‘s, review of relevant sources of information and observing when traffic officer generate decision on the road traffic accident, control mechanisms and an implementation. Relevant information is implied as record of accident data collected in hard copy format. Hard copy format of data is translated into spreadsheet format in excel. The record collected was 1863 and most of the instances are concluded with similar values of attribute and needs to be preprocessed using WEKA data mining technique before stored into relational database. The filtered data then exported to MySQL database after saved in ARF file format and converted back to CSV file format.

The objective of knowledge acquisition is to gather the required knowledge, interpreting the acquired knowledge, analyzing and validating the knowledge content. Based on the acquired knowledge, the KBS was designed using decision tree model. Therefore, knowledge acquisition process of this thesis was based on domain expert interviewing, observing and reviewing of related documents.

The researcher acquires knowledge using two types of knowledge acquisition methods which are manual and automatic knowledge extraction.

3.2.1. Interviewing Domain Experts

Knowledge was acquired and collected from human experts in the domain area at traffic police office of two stations. To gather the required knowledge semi-structured interview technique was used. Since one of the main focuses of this research is eliciting relevant knowledge from the domain experts and make relevant decision interview must important to identify problem and determine the future solution. There are four (4) domain experts, two (2) in station1 and another two (2) in station2. All experts have equal knowledge, skill and perception on road traffic accident. So the researcher selected two experts purposely, one from station1 and another one from station2 using purposive sampling technique. These experts were selected due to their managerial position and experience in managing road traffic accident. This sample was taken due to limited number of experts in the managerial position of the Jimma traffic office. During the interview phase, the main challenge was willingness of domain expert to share their expertise and experience, which was important to acquire the relevant knowledge. The interview with expert covered issues such as how the experts diagnose road traffic accident, what are the techniques used to identify the pillar symptoms of the problem, the procedures of diagnosing and what are the possible solutions recommended to the problems. During the extensive discussion, the researcher acquired the relevant knowledge which was significant to generate the rules. In addition, the domain experts were actively participated throughout the research work and they were consulted to confirm the correctness of the acquired knowledge. During face to face communication, the acquired knowledge from domain experts was recorded manually by using pen and paper sheet. Profiles of traffic police domain experts participated in the interview process were illustrated in the Table 3.1 below:

Table 3.1 Domain Expert Profiles

No	**Traffic based Educational Position and qualification**	**Qualification**	**Position**	**Working Area/zone**

1.	Inspector	BA Degree	Traffic police manager	Number One
2.	Inspector	BA Degree	Traffic police manager	Number Two
3.	Sajin	BA Degree	Assistant manager	Number One
4.	Sajin	BA Degree	Assistant manager	Number Two

As indicated in the Table 3.1 above, the traffic police expert were devoted to providing decision making service on road traffic accident based on the data of weakly/monthly recorded accident data and preventive idea from the traffic policemen were on duty.

The knowledge acquired from traffic expert by semi structured interview questions and result was discussed below:

Traffic problems: The first and basic question was "What are adapted traffic problems (accident and congestion) in Jimma city?" Because the researcher limited the conduct with the accidents such as vehicle to vehicle and vehicle crash pedestrian. Other type of accident did not included in this research. The answer of traffic officer was there are various problems that are occurred due to road traffic accident in Jimma town as high severe such as disability, loss of property and fatality. Highly severe problems occurred are up to 15-20 per month and other less sever accident problem are not counted.

Case of accident: The second question was "What are the main risk factors that cause an accident repeatedly?" This is important question for the making the right decision and make preventive measurement. The major risk factors or accident probable causes are driver, pedestrian, vehicle, environment and roads. The main risk factors of the accident nowadays in Jimma town are young age driver and driving with maximum speed in town at where minimum number of traffic policemen. As interviewee answer question road infrastructure is another risk factor. The following components other risk factors:

A. **Driver (driving speed, use alcohol and health):** is at the core of the road traffic injury problem. Speed influences both crash risk and crash consequence. The physical layout of the road and its surroundings can both encourage and discourage speed. Crash risk increases as speed increases, especially at road junctions and while overtaking as road users underestimate the speed and overestimate the distance of an approaching vehicle.

In addition impairment by alcohol is an important factor influencing both the risk of a road crash as well as the severity of the injuries that result from crashes. When driver ride long time and distance the same problem related to his fatigue condition also happened. Lack of skill means driver has no license sometimes victimized to the accident. The frequency of drinking and driving varies between countries but it is almost universally a major risk factor for road traffic crashes.

Fatigue or sleepiness is associated with a range of factors. Some of these factors with relevance to road traffic are long distance driving, sleep deprivation and the disruption of circadian rhythms.

B. Road condition: Road states three different factor as the traffic expert argued it. The first is road surface (asphalt or ground), road surf (road is dry or covered or not with wet and muddy), road orientation (road have or not traffic policemen, traffic signal).These are the most important factor that help traffic officer to make decision after accident. Most of area in Jimma there are no traffic signal that assist driver and pedestrian.

C. Pedestrian condition: When the accident I between pedestrian and vehicle pedestrian can be victimized to accident with lack of awareness how to use road side walk, when and where to cross the road, without caring use the road.

D. Mechanical problem: When the cause of accident is because of vehicle disappointment and not either driver, road or pedestrian. This is not included in this study as a factor that help traffic officer to make decision.

Predetermined Policies to represent above cases: The third question was "Are there any predetermined policies (rule and standard) to manage an accident based on their severe?" For be sure in his/her decision manually an expert must have any predetermined policies that help for making decision and take preventive measurement to improve accident controlling mechanisms. Traffic officer responded to researcher as there are no regular policy (rule and standard) that help for transferring the future solution on report. But based on the area on accident happened and situation recorded the following rule performed to achieve on right decision.

When there were accidents vehicle crash pedestrian, kebele …., particular area …., traffic sector around….., Driver speed was ….., Road signals ….., Traffic police was ….., Traffic congestion

….., injured pedestrian road using condition…..., **Then** Traffic officer transfer future direction and possible alternative solution.

When there were accident that vehicle crash vehicle, kebele …., particular area …., traffic sector around….., Driver1 speed …., Driver2 speed …., Road signals ….., Traffic police was ….., Traffic congestion ….., Driver1 road using condition was ….., Driver2 road using condition….., **Then** Traffic officer transfer future direction and possible alternative solution.

Therefore the traffic officer held meeting with traffic policemen on duty specifically where accident are happened frequently and discuss on what future solution should be. Then all traffic police men on duty of accident area have discussion on it and implement the final solution.

Preventive measurement: The fourth question was "How to determine the probable risk of an accident in future time and what preventive measurements were considered?" There are kinds of preventive measurements that are taken by traffic police expert to handle an accident. To identify this it is important to raising this question. As traffic officer answer this question a situation of accident, how the environment, on what road side accident happened (example, if the car crash pedestrian technically on what road side pedestrian is crashed like information will gathered). If the road infrastructure is poor and counted as the risk, this will considered as preventive measurement. If there is no any signal, this also considered as preventive measurement.

Decision making process: The Fifth question is "What are the main decisions the traffic officer make to solve road traffic problem?" For this kind of accident is that necessary to having measurement or not is the decision. Because sometimes some types of accidents not needs the measure because their probability to be happened in future is rare. The decision transferred by traffic officer include installation of signal, increase number of traffic policemen, request for road maintenance, punishment, build awareness of community how to use road, etc.

Guideline for decision: The Sixth question was "Do you have a standard guideline that you use for the decision making and approval process?" To make decision a traffic police must have a guideline. Because there are 3-15 accident happened per a day in Jimma Town. So it is difficult to make decision without any guideline.

The seventh question was “When, where decision making idea was generated or incase what influential matter?” This also important question because what factors influence the expert to make a decision after an accident? May the accident happen repeatedly at that area, and there is another special reason.

The eighth question is “What are the common decision on the accident and congestion?” Since there is traffic accidents and congestions for both it needs a decision that solves the problem of both at where they are available.

Challenges in traffic accident control: The last and ninth question is “How you manage the challenges during accident problem and making decision to improve road traffic control?” There are challenges in accident problem solving and implementation of the solution. So it is important to know that challenges. The challenge is after an accident happened it takes long time to transfer preventive measurement, even if the decision is transferred the implementation takes a long period of time.

3.2.2. Knowledge Acquired from Data Mining

Knowledge acquisition is a process of identifying the knowledge, representing the knowledge in a proper format, structuring the knowledge, and transferring the knowledge to a machine. This process can be affected by the roles of the knowledge engineer, the expert and the end user (Bhambr, 2011).

Traditional knowledge acquisition techniques including on site observation, protocol analysis, structured and semi structured interview and others will be used. As stated by Cornelius (2005), there are significant problems with each of these techniques. None of them guarantees consistency and integrity in the knowledge base. Some of the problems mentioned are: labor intensive, expensive to implement, expert conservatism and unwarranted biases.

Due to the aforementioned problems knowledge engineers look for other means to expand rule set and verify the rules already in the knowledge base. As a result Mihaela (2006) and Charles & Duminda (2002) stressed the need for developing automated techniques for knowledge acquisition.

In addition, knowledge can also be acquired from large collection of dataset by using knowledge discovery tools. This type of knowledge is called hidden knowledge.

The researcher also used data mining approaches to acquire knowledge for the purpose of matching similarity of knowledge representation result from interview with traffic expert and recorded knowledge representation.

Considering the limitations mentioned above for acquiring knowledge from experts using traditional knowledge acquisition techniques, the researcher used data mining techniques for the development of the knowledge based expert system. Hence data mining specially classification algorithms are employed to generate cases (knowledge) for the knowledge base.

In Jimma traffic office specific data of road traffic accident was recorded manually with standardized format on paper and difficult of using making decision on road traffic accident in real time. Traffic police officer also prepares report from this manually recorded accident data on Microsoft word and spreadsheet.

With the use of data mining techniques it is possible to extract interesting and useful knowledge and this knowledge can be used by experts for efficient and enhanced decision making process.

In addition, Knowledge acquisition is a complex and time consuming stage during rule based system development (Medi, 2008). For knowledge generation and model building, classifier algorithms such as J48 and PART are employed and their result is compared to generate best rules and representative model for the knowledge based expert system. Knowledge representation schemes such as frames, cases, semantic rules, and rule-based systems exist to represent knowledge (Charles & Duminda, 2002).

For this study the data is collected from traffic police office of Jimma town from 2007 to 04/2009 Ethiopian calendar. The researcher used data (cause) that encourages traffic police officer to make decision that improves road traffic accident controlling mechanism and managing the data for the future use. The data which are encouraging traffic police officer to make preventive decision after accidents are classified based on the severity of an accident. In this research three decisive classes are contributed to achieve on the decision made by the road

traffic officer. These classes are extreme serious (severe), highly serious (highly severe) and Less severe.

High Serious (severe): When an accident from both vehicle to vehicle crash and vehicle crash pedestrian is highly serious in terms of road problem, driver problem and other environmental is high severe. As a result road surface is either asphalt or ground; road surf is either muddy or wet, road is oriented as the straight plain or signals, congestion of people and crowds of car, etc are occurred it is high severe.

Less serious (severe): When there are low volume of an accident from both vehicle to vehicle crash and vehicle crashes pedestrian are occurred it is less serious (severe). Most of the accident causes are not contributed in the accident.

The following table 3.1 summarizes patterns of accident that are the number of attributes and number of record that researcher has collected from the road traffic office of Jimma Town. The datasets collected from traffic office within different years have the same number of attributes since they are prepared for the sake of reports to Oromia road traffic commission with similar report formats.

Table 3.2 Numbers of Attributes and Record

Number of attributes	**Amount record**	**Recorded year**	**Station/Zone**
31	173	2013	No1
	250	2013	No 2
31	230	2014	No1
	241	2014	No 2
31	220	2015	No1
	148	2015	No 2
31	210	2016	No1
	225	2016	No 2
31	94	2017 up to February	No1
	72	2017 up to February	No 2
Total	1863		

As a researcher discussed with the traffic expert, from this 31 attributes some attributes could represented group of attributes for example: for attribute that is ‘particular place’ can represents attributes such as Zone, city, kebele and particular place. Another good example is the cause that ‘pedestrian’ there are attributes not required for the system such as injured pedestrian name, sex, age, etc. As a traffic expert argued, these kinds of attributes are picked only for additional information and reported to the court for lawyer.

Total amount of record researcher was converted from the hard copy to the soft copy dataset are 1863 and researcher categorized them based on the type of road traffic accident that are vehicle crash vehicle and vehicle crash pedestrian. The number of record was not only 1863 but most of data recorded are similar with this amount of dataset. So the researcher jump over similar data found in records and even filter them using WEKA data mining technique using remove duplicate instance under unsupervised algorithm.

3.2.2.1. Data Preprocessing

Today’s real-world databases are highly susceptible to noisy, missing, and inconsistent data due to their huge size (often several gigabytes or more) and origin from multiple, heterogeneous sources (Zhuang et al., 2009). Therefore, prior to giving the data to a data mining tool, preprocessing of the data is necessary. Preprocessing the data includes multiple steps to assure the highest possible data quality, thus efforts are made to detect and remove errors, resolve data redundancies, and taking into account of the patient privacy, to remove patient identifiers (Inderpal, 2013). Data processing techniques, when applied before mining, can substantially improve the overall quality of the patterns mined and/or the time required for the actual mining (Zhuang, et al., 2009). There are a number of data preprocessing techniques. Data cleaning can be applied to remove noise and correct inconsistencies in the data. Data integration merges data from multiple sources into a coherent data store. Data transformations, such as normalization, may improve the accuracy and efficiency of mining algorithms involving distance measurements. Data reduction can reduce the data size by aggregating, eliminating redundant features. These techniques are not mutually exclusive; they may work together for a better data quality. Data mining requires access to data. The data may be represented as volumes of records

in several database files or the data may contain only a few hundred records in a single file. According to (Inderpal, 2013), there are three common ways to access data for data mining:

i. Data can be accessed from a data warehouse,

ii. Data can be accessed from a database

iii. Data can be accessed from a flat file or spreadsheet. In our case we used dataset from a flat file or a spreadsheet.

In this case the researcher acquired the data from the hard copy (paper). Record on the paper format was converted dataset. The researcher used dataset from a flat file or a spreadsheet. The researcher process road traffic accident data acquired and 1863 instance stored in excel and converted into CSV file format. From this 1863 vehicle crash vehicle data was 529. In addition 1324 data are vehicle crashes pedestrian data. 873 out 1324 instance are replicated data of 451 original pure data. Using WEKA and its experimental tools replicated data removed and 59 instance values are stored to knowledge base and imported to MySQL database for retrieval. In addition vehicle crash pedestrian replicated data are removed and 451pure instance values are imported to MySQL database which act as knowledge base for the collection, representation and transformation of knowledge.

3.2.2.2. Data formatting

Dataset acquired from road traffic office of Jimma from both station and founded with hard copy with the required format. The researcher prepared this hard copy format of data in order and converted to the soft copy in micro soft office excels. Dataset accepted with Excel spreadsheet has 18 attributes that accidents are recorded as instance.

These attributes are accident moment, road surface, road surf, road orientation, congestion, driver1health condition, driver1potential drive, driver one drive side, driver two health condition, driver two potential drive, driver two drive side, weather, pedestrian health condition, pedestrian road side, severity, future direction and solution.

Instances of the Dataset written in Excel spreadsheet are converted to CSV file format and saved with CSV Comma Delimited. An instance shown in figure 3.3.below

```
@relation vehped

@attribute type {'vehicle to pedestrian'}
@attribute moment {day,night}
@attribute parea {'agaro bar','ajip bajaj tara','ajip mikael',yetebaberu,'seka bar','kulo bar','merkato taxi tara','merkato bajaj tara',kera,'bacho bore','ajip kolo bar','ferenji arada
@attribute surface {asphalt,ground}
@attribute surf {wet,dry,mud,muddy}
@attribute orient {'no signal','no traffic','signal and traffic'}
@attribute congest {'crowd of people','crowd of car','average both','minimum car and people','crowd of Car'}
@attribute dlhealth {normal,'no liscence',drank,'physical and psychological'}
@attribute dlpot {'excessive speed','average speed',roundabout,'no lisence'}
@attribute dlside {'right side','wrong side','cross road',roundabout,rounabout}
@attribute pedhealth {normal}
@attribute pedside {'cross road',standing,'right side','wrong side'}
@attribute weather {sunny,rainy,cloudy,'cloudy ',Sunny,Rainy}
@attribute severity {'high severe','less severe'}
@attribute pmeasurment {yes,no}
@attribute fsolution {'Build wet pavement at road side and Install signal(direction)','Install barrier for pedestrian (zebra) and driver limit speed','Install barrier (Zebra) for pedst

@data
'vehicle to pedestrian',day,'agaro bar',asphalt,wet,'no signal','crowd of people',normal,'excessive speed','right side',normal,'cross road',sunny,'high severe',yes,'Build wet pavement
'vehicle to pedestrian',day,'ajip bajaj tara',asphalt,dry,'no signal','crowd of people',normal,'excessive speed','wrong side',normal,standing,sunny,'high severe',yes,'Install barrier f
'vehicle to pedestrian',day,'ajip mikael',asphalt,dry,'no signal','crowd of people',normal,'excessive speed','cross road',normal,'right side',sunny,'high severe',yes,'Install barrier (
'vehicle to pedestrian',day,yetebaberu,asphalt,dry,'no signal','crowd of people',normal,'excessive speed',roundabout,normal,'wrong side',sunny,'high severe',yes,'Install barrier (Zebra
'vehicle to pedestrian',day,'seka bar',asphalt,dry,'no signal','crowd of people',normal,'average speed','right side',normal,'cross road',sunny,'high severe',yes,'Install barrier (Zebra
'vehicle to pedestrian',day,'seka bar',asphalt,dry,'no signal','crowd of people',normal,'average speed','wrong side',normal,standing,sunny,'high severe',yes,'Install barrier (Zebra) fo
'vehicle to pedestrian',day,'seka bar',asphalt,dry,'no signal','crowd of people',normal,'average speed','cross road',normal,'right side',rainy,'high severe',yes,'Install barrier (Zebra
'vehicle to pedestrian',day,'kulo bar',asphalt,dry,'no signal','crowd of people',normal,'average speed',roundabout,normal,'wrong side',rainy,'high severe',yes,'Install barrier (Zebra)
'vehicle to pedestrian',day,yetebaberu,asphalt,dry,'no signal','crowd of people',normal,roundabout,'right side',normal,'cross road',cloudy,'high severe',yes,'Install barrier (Zebra) fo
'vehicle to pedestrian',day,'merkato taxi tara',asphalt,dry,'no signal','crowd of people',normal,roundabout,'wrong side',normal,standing,cloudy,'high severe',yes,'Install barrier (Zebr
'vehicle to pedestrian',day,yetebaberu,asphalt,dry,'no signal','crowd of people',normal,roundabout,'cross road',normal,'right side',cloudy,'high severe',yes,'Install barrier (Zebra) fo
'vehicle to pedestrian',day,yetebaberu,asphalt,dry,'no signal','crowd of people',normal,roundabout,roundabout,normal,'wrong side',cloudy,'high severe',yes,'Install barrier (Zebra) for
'vehicle to pedestrian',day,yetebaberu,asphalt,dry,'no traffic','crowd of people',normal,'excessive speed','right side',normal,'cross road',sunny,'high severe',yes,'Build warning sign
```

Figure 3.1 Test instances of vehicle to vehicle crash without any experiment in WEKA

Dataset stored in WEKA saved with file format .ARFF and researcher applied and test with data mining modeling experiments.

Like any other software, WEKA needs data to be prepared in some formats and file types. The datasets provided to this software were prepared in a format that is acceptable for WEKA software. WEKA accepts records which attribute values are separated by commas and saved in an ARFF (Attribute-Relation File Format) file format (a file name with an extension of ARFF i.e. FileName.arff). At first the integrated dataset was in an excel file format. To feed the final dataset into the WEKA DM software the file is changed into other file format. The excel file was first changed into a comma delimited (CSV) file format. After changing the dataset into a CSV format the next step was opening the file with the WEKA DM software.

These data imported to MYSQL database with apache server configuration using C language written on PHP. The data formatted in all these recycle will be displayed o web user interface.

3.2.2.3. Dataset Description

Attributes for the database table are not the same as number of attribute specified in the table 3.1. For this study the researcher agreed on the current situation of the traffic accident data management. Currently the traffic police officers are using about 31 attributes for the road traffic data management and control on manual or paper format. The collected data from the paper collection have a total of 31 attributes. From these collected accident record from paper 12 attributes was not considered as independent attributes and some of them can represented with single attribute. Also these 12 attributes have no operational contribution in generation of road traffic accident control mechanism and cannot used as a key factor in decision making process. In this study 19 attributes are selected considering decision making process and prepare preventive measure. The table 3.2 below shows the data set description of road traffic accident data.

Table 3.3 Data set and Database description

No	Attribute	Description
1.	Type of accident	Type of crash or accident
2.	Accident moment	Whether accident happened in a moment of day or night
3.	Particular area	Area an accident happened
4.	Road Surface	Whether the road surface is asphalt or ground.
5.	Road Surf	Whether the road surface is dry, muddy, or wet.
6.	Road Orientation	How the road is oriented including signal or straight plain
7.	Congestion	Whether the road is covered with number of people or Vehicle
8.	Driver One Health Condition	Whether driver have physical problem, fatigue, drank or normal
9.	Driver One driving Potential	Whether driver drive with excessive speed, without skill, license, or no problem
10.	Driving One Road Side	Whether driver driving on wrong side, right side and crossing
11.	Driver Two Health Condition	Whether driver have physical problem, fatigue, drank or normal
12.	Driver Two Driving Potential	Whether driver drive with excessive speed, without skill, license, or no problem
13.	Driver Two Driving Road	Whether driver driving on wrong side, right side and crossing

	Side	
14.	Weather Condition	The weather condition
15.	Pedestrian Health condition	Whether pedestrian have physical problem, and drank or normal
16.	Pedestrian Walking Side	Whether pedestrian is walking on wrong side, right side and crossing
17.	Accident Severity	The severity of the accident
18.	Measurement	Decision made and Measurement taken on accident
19.	Solution	Final solution

3.2.2.4. Attribute Selection

To select the best attributes for data mining, the researcher uses information gain method which exists in WEKA data mining tool. The attribute with the highest information gain is selected as the splitting attribute. This attribute minimizes the information needed to classify the instances in the resulting partitions and reflects the least impurity in these partitions. Entropy (impurity) is used to measure the information content of the attributes. High entropy means the attribute is from a uniform distribution where as low entropy means the attribute is from a varied distribution.

Before calculating the information gain of the attributes the researcher discussed with domain experts to select the most significant attributes for decision making. According to road traffic expert all attributes are important to reach on the right decision for the accident identifying the accident severity. If the accident severity is extreme, the preventive measurement or the best alternative solution would be taken urgent. If the accident severity is less severe, the alternative solution might not necessary.

For the achievement of objective, 19 attribute selected by researcher and traffic expert are description as case attributes regarding name and data type in table 3.3.

Table 3.4 selected attribute for Input Case descriptions

No	**Attribute**	**Data Type**
1.	Type of accident	VARCHAR
2.	Particular Area	VARCHAR
3.	Moment accident happened	VARCHAR
4.	Road Surface	VARCHAR

5.	Road Surf	VARCHAR
6.	Road Orientation	VARCHAR
7.	Congestion	VARCHAR
8.	Driver One Health Condition	VARCHAR
9.	Driver One driving Potential	VARCHAR
10.	Driving One Road Side	VARCHAR
11.	Driver Two Health Condition	VARCHAR
12.	Driver Two Driving Potential	VARCHAR
13.	Driver Two Driving Road Side	VARCHAR
14.	Weather Condition	VARCHAR
15.	Pedestrian Health condition	VARCHAR
16.	Pedestrian Walking Side	VARCHAR
17.	Accident Severity	VARCHAR
Attribute for the road traffic accident solution		
18.	Decision/Measurement	Text
19.	Solution	Text

Some attribute that are included in accident record in paper format are removed. This is determined by expert based on their role and quality for the decision making process. Since the developed KBES system is limited by the final solution that is accident preventive or reducing measure, an attributes that play role in solution and decision making only selected. Example, for zone, city, woreda and kebele it is enough representing attribute "particular place" since site study was Jimma town. These removed attributes are listed in table 3.4 bellow.

Table 3.5 Removed attribute from original data

No	**Attribute**	**Data Type**
1.	Zone	VARCHAR
2.	City	VARCHAR
3.	Kebele	VARCHAR
4.	Vehicle table code	VARCHAR
5.	Name of driver	VARCHAR
6.	Driver sex	VARCHAR

7.	Driver age	INTIGER
8.	Name of vehicle	VARCHAR
9.	Type of vehicle	VARCHAR
10.	Injured person name	VARCHAR
11.	Injured person sex	VARCHAR
12.	Injured person age	INTIGER

As table 3.5 shows the researcher take time to discuss with traffic expert to decide add and remove attribute that are serving as independent attribute and play major role in decision making process by traffic officer. As a result traffic expert improve these attributes are not serve for decision making process the researcher minimize Eleven (12) attribute from 31attributes. The remained19 attribute which 17 attributes are involved in vehicle to vehicle crash and 16 for vehicle to pedestrian crash with three and two unique attributes respectively.

3.2.2.5. Data Mining for data representation and modeling

The objective of this step is to apply preprocessing technique algorithms on Road traffic accident data set which have been collected from road traffic office and develop a model that can perform filtering of best attribute and its value (instance) for Knowledge based expert system development. Basically supervised and unsupervised filtering algorithms are available in WEKA software of data mining.

A total of three experiments aiming at building predictive models are undertaken. The sampled data set contains 1863 instances. The data set contains 19 attributes and all of them are involved in all experiments. In addition after undertaking a number of experiments, default value of parameters is taken into consideration for each classifier algorithm since it allows achieving better accuracy compared to modifying the default parameters values. The researcher conducted three experiments for three classification algorithms: namely J48 pruned and PART.

Experiment #1 using J48 Pruned for knowledge modeling

Decision tree is a graphical representation of the relations that exist between the data in the database. It is used for data classification. The result is displayed as a tree, hence the name of this

technique. Decision trees are mainly used in the classification and prediction. It is a simple and a powerful way of representing knowledge.

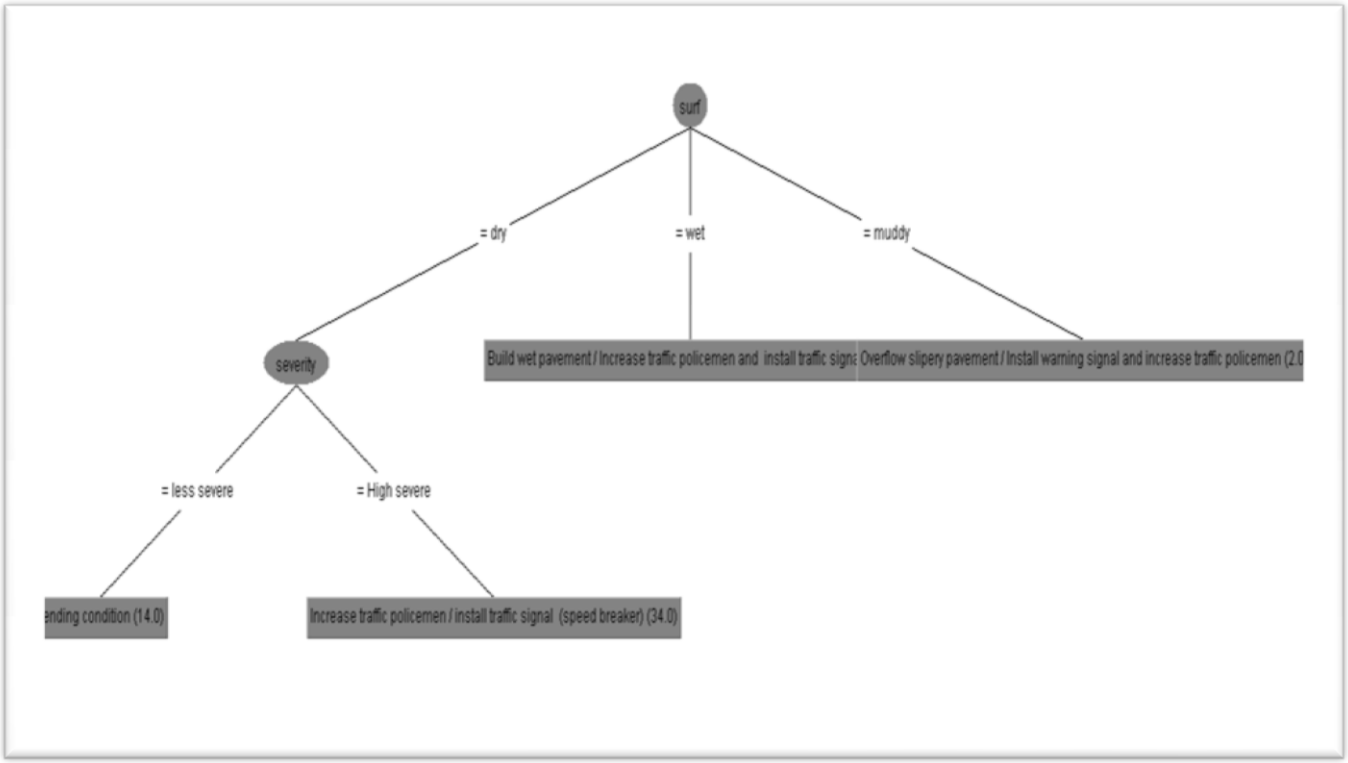

Figure 3.2 Sample J48 decision tree

These experiment help the researcher to select high frequency of case encourage extremely high severe that make traffic officer to generate future solution with future direction frequency greater than 90%. Evaluation test for J48 classification algorithm, the researcher used cross percentage split of 40%. Here all 40% of 1863 instances were tested in WEKA. The following results are generated as illustrated in table 3.6 below.

Table 3.6 J48 classification percentage split (40%) result

S. No	**Evaluation on test split on cross validation fold 10**	**Result**
	Correctly Classified Instances	88.5965 %
	Incorrectly Classified Instances	11.4035 %
	Time taken to build model	0.08 seconds
	Total Number of Instances	456

	Ignored Class Unknown Instances	9
	Kappa statistic	0.7992
	Mean absolute error	0.0356
	Root mean squared error	0.1701
	Relative absolute error	20.2224 %
	Root relative squared error	56.477 %
	Coverage of cases (0.95 level)	91.1111 %
	Mean rel. region size (0.95 level)	13.5802 %

In this selected experiment of WEKA for classification algorithms are used for the purpose of making error free of instance which are written by researcher before imported to MYSQL database, process data selection, determine future direction solution (what will be the solution and for what cases solution will determined).

Experiment #2 using PART rule induction for knowledge representation

PART is a rule-based classifier uses a set of IF-THEN rules for classification. An IF-THEN rule is an expression of the form IF condition THEN conclusion. The "IF"-part (or left-hand side) of a rule is known as the rule antecedent or pre condition. The "THEN" part (or right-hand side) is the rule consequent (Chen 2009).

This experiment conducted under percentage splits technique using 90 % of instances for training and the remaining for testing with default parameters of WEKA and the algorithm generates a model with 44 rules. As shown in table 3.8 Correctly Classified Instances are 1615 which means 93.5556 % and Incorrectly Classified Instances are 248 which are 6.4444 % from Total Number of Instances of 1863.

Table 3.7 sample of Rules extracted by PART Classification algorithm

surf = dry AND
pedside = standing AND
d1pot = excessive speed AND

d1side = wrong side
severity = high severe
orient = no signal: Increase traffic policemen / install traffic signal (zebra) and train people awareness on road using (16.0)

surf = wet AND
orient = no signal and no traffic AND
d1pot = excessive speed AND
severity = high severe
pedside = standing AND
congest = crowd of car: Build wet pavement / Increase traffic policemen and install traffic signal (zebra) (8.0)

orient = no signal and no traffic AND
d1pot = excessive speed AND
severity = high severe
pedside = cross road AND
congest = crowd of people: Install traffic signal, speed breaker and increase trafic policemen (8.0)

orient = no signal and no traffic AND
severity = high
pedside = cross road AND
d1pot = excessive speed: Install speed breaker and traffic signals (3.0)

pedside = wrong side AND
severity = high severe
orient = no signal AND
d1pot = average speed AND
congest = crowd of people: Train people awareness on road using and Install signal (23.0)

surf = muddy AND
orient = no signal AND
severity = high severe
pedside = wrong side: Overflow slipery pavement / Install warning signal and train people awerness on road using (11.0)

surf = muddy AND
severity = high severe
congest = crowd of car: Overflow slipery pavement / Install warning signal (zebra) and increase traffic policemen (25.0)

Other rule generated using PART classification algorithm are illustrate in the appendix IV. While PART classification run on the test option of cross validation fold ten (10) for 1614 instance of 16 attributes. The result illustrated in table 3.7 below shown on 86.6228% of 1688 Correctly Classified Instances and 175 instance of Incorrectly Classified was 13.3772 %. Here the number of rule produced with PART classification algorithm on percentage split of 40% was 44.

Table 3.8 Confusion Matrix for PART rule induction

Confusion Matrix				
A	**b**	**C**	**d**	**Classified as:**
14	0	0	0	A= Pending condition
0	9	0	0	B=Build wet pavement / Increase traffic policemen and install traffic signal
0	0	34	0	C= Increase traffic policemen / install traffic signal (speed breaker)
0	0	0	2	D = Overflow slippery pavement / Install warning signal and increase traffic policemen

3.3. Conceptual Modeling

Once the required knowledge is acquired from road traffic accident record, traffic experts and relevant documents by data mining, the second step for development of KBES prototype is modeling the knowledge acquired from both source.

Conceptual Modeling of domain knowledge implies capturing the static structure of information and knowledge types (Lidtke, 2003). Decision trees (DTs) are modeling tools that are used in a variety of different settings to organize and break down clusters of data (Scott, 2004). The knowledge modeling step involves organizing and structuring of the knowledge gathered during knowledge acquisition. This activity provides an implementation independent specification of the knowledge to be represented in the knowledge base. Knowledge modeling is the concept of representing information and the logic for the purpose of capturing, sharing and processing knowledge to simulate intelligence (Makhfi, 2011). Here, the basic concepts that tell the main activities and decisions are made to solve cases in the domain are modeled. Conceptual modeling is a crucial step in the knowledge acquisition process so as to understand well the problem domain and to prepare the knowledge representation phase.

There are different conceptual modeling techniques that could be determined by data mining classification algorithm.

For this study decision tree structure was determined by WEKA data mining discovered by J48 classification algorithm. As Richard (2000) introduced decision tree is used to model how road traffic accident management and control process. Because it performed since one of the most extensively applied methods of conceptual modeling is decision tree.

3.3.1. Decision Tree Structure and Logical View of RTACKBES

Representing RTA management and control with decision tree model help the traffic police officer to understand accident problems, to reach on the right decision such as prediction of severity of accident, taken preventive measurement and best alternative solution to improve the controlling mechanisms. Therefore, the system improves the road traffic accident problem causes diagnosing service; minimize the burdens of traffic police in diagnosing and exploring solution for the accident, fill the gaps of limited manpower, delay of time and backlog.

The common causes of each problem are used to determine the presence of specific cases that has been illustrated in figure 3.2 using decision tree structures. Based on the decision tree structures, some problems which have similar characteristic in nature were incorporated in the knowledge base to show that some causes can also be causes for other problems. Therefore, similar procedures have been followed for all problems occurring in the computer system. Some of the causes were modeled by using decision tree as follows:

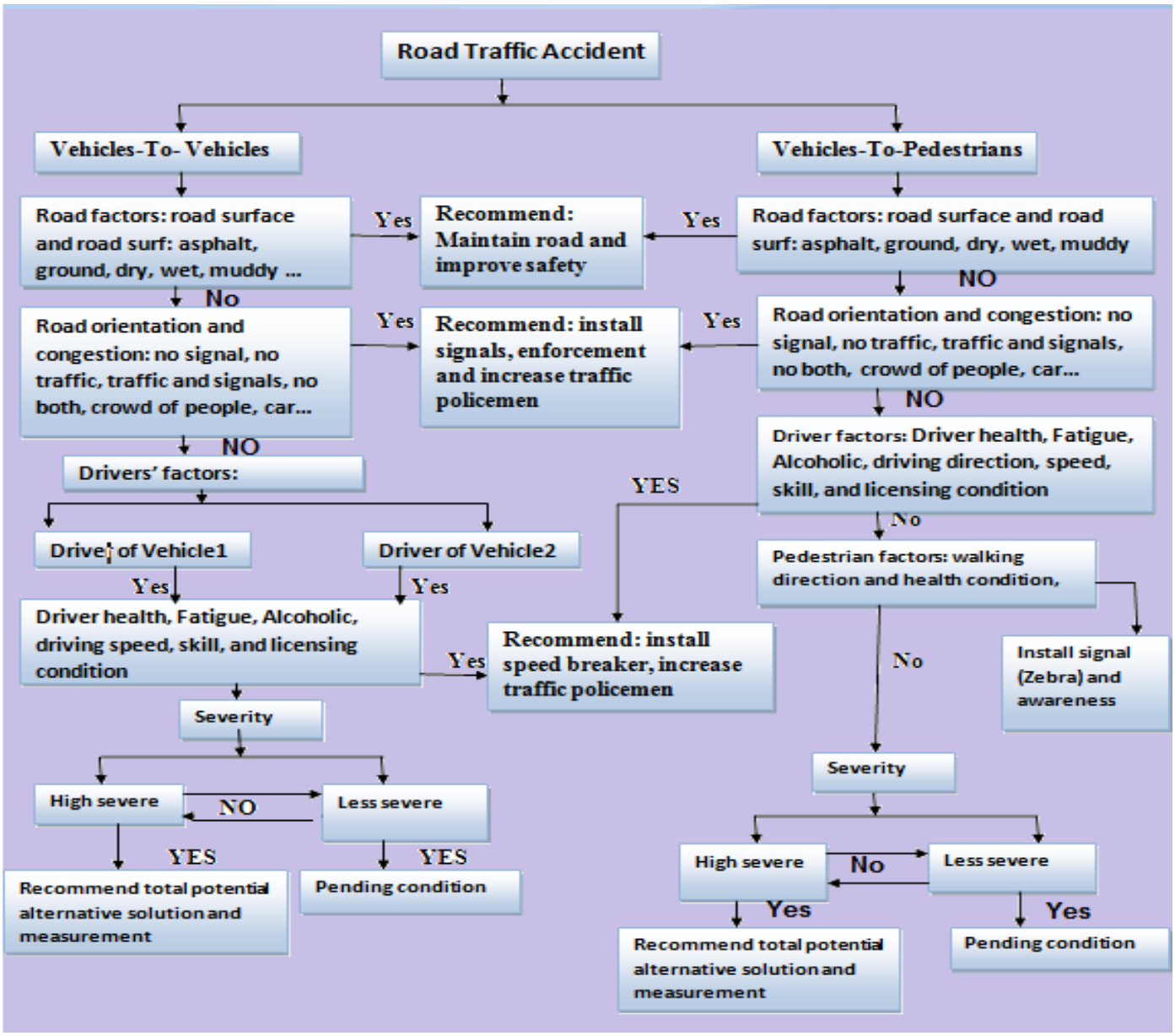

Figure 3.3 Decision Tree for road traffic accident control

As illustrated in the above figure 3.3 the road traffic accident control process started from the icon road traffic accident. The type of road traffic accident the researcher focus on was vehicle to vehicle crash and vehicle to pedestrian crash. These two accidents risk factors play major role in road traffic are road, driver and pedestrian. For the problem occurred the solution are recommended by the system is also involved.

3.4. Knowledge Representation

Knowledge representation is third steps in the process of knowledge based system development. Knowledge representation is the process of interpreting acquired domain knowledge into computer understandable form by using knowledge representation methods. The researcher represented knowledge with rule based representation of cases collected by human expert and processed manually. The acquired domain knowledge is represented as a set of -IF -THEN- rules in the prototype. The -if- side (also known as the left hand side) of the equation which states the condition(s) that must be true in order to fire the rule and the -then- side (right hand side) of the equation specifies the appropriate action to be taken. The inference engine evaluates the -if- portion of a statement and concludes whether a goal is satisfied or not. If the goal is not satisfied, the inference engine proceeds to the next rule until the conditions are satisfied. A rule is a conditional statement that links the given conditions to actions. Rules in the knowledge base were constructed based on the decision tree structure on conceptual model discussed above. To make easy and understandable web development PHP rules, the acquired knowledge from the domain experts represented using the -IF-THEN- form. The rules are the base for the construction of KBS. The following are sample rules which are incorporated in the knowledge base.

Researcher used the rule representation for the system and the system developed is properly illustrating this rule. The following rule is represented the knowledge in the knowledge based expert system. These sets of rule are extracted from data mining application experiment PART and J48 algorithm result described n section 3.1.1.2.5 above. Selected attributes are used in report form of road traffic accident reported to the traffic police head office.

If accident Type is vehicle to vehicle crash

AND Road surface is Asphalt

AND Accident moment is day time

AND Road Surf Muddy

AND Road is not oriented by Signal

AND Congestion with number of people

AND Driver one has no physical problem OR Not Drank OR Not Fatigue

AND Driver two has no physical problem OR Not Drank OR Not Fatigue

AND Driver one driving with average speed

AND Driver Two driving with average speed

AND Driver one driving on Wrong direction

AND Driver Two driving on Right direction

AND Weather condition is Rainy

THEN

Preventive measurement needs:

Yes

Final Solution:

Accident Injure severity is highly severe

Install a signal that a driver be care on both directions

Built the way for the flow of wet

If accident Type is vehicle to vehicle crash

AND Road surface is Asphalt OR Ground

AND Road Surf is dry

AND Accident moment is day time

AND Road is not oriented by Signal

AND Congestion with number of People and Car

AND Driver one has no physical problem OR Not Drank OR Not Fatigue OR Normal

AND Driver two has no physical problem OR Not Drank OR Not Fatigue OR Normal

AND Driver one driving with excessive speed

AND Driver Two driving with average speed

AND Driver one driving on Wrong direction

AND Driver Two driving on crossing road

AND Weather condition is dry

AND Accident Injure is highly severe

THEN

Preventive measurement needs:

Yes

Final Solution:

Accident Injure severity is highly severe

Install traffic sign to assist driver to reduce speed

Increase number of traffic police men

If accident Type is vehicle to vehicle crash

AND Road surface is Asphalt OR Ground

AND Road Surf is dry

AND Accident moment is day

AND Road is oriented by Signal

AND Congestion with number of People and Car

AND Driver one has No license OR Normal

AND Driver two has no physical problem OR Not Drank OR Not Fatigue OR Normal

AND Driver one driving with excessive speed

AND Driver Two driving with average speed

AND Driver one driving on Wrong direction

AND Driver Two driving on his own road side

AND Weather condition is normal

AND Accident Injure is highly severe

THEN

Preventive measurement needs:

Yes

Final Solution:

Accident Injure severity is highly severe

Increase number of traffic police men

If accident Type is vehicle crash pedestrian (person)

AND Road surface is Asphalt OR Ground

AND Road Surf is dry

AND Accident moment is day

AND Road is oriented by Signal

AND Congestion with number of People and Car

AND Driver has license OR skill enough

AND Driver has no physical problem OR Not Drank OR Not Fatigue OR Normal

AND Driver driving with excessive speed

AND Driver is driving on his own road side

AND Weather condition is normal

AND Pedestrian has no physical problem OR Not Drank OR Normal

AND Pedestrian is crossing the road

AND Accident Injure is highly severe

THEN

Preventive measurement needs:

Yes

Final Solution:

Accident Injure severity is highly severe

Increase number of traffic police men

Build warning sign (Zebra sign)

Train people Awareness and have knowledge in road using during walk

Similar procedures have been used for all rules incorporated in the knowledge base. To summarize the chapter, the transformation process of the acquired knowledge from human expert, document analysis, observation and data mining experiment is helpful to investigate the specific attributes which are significant in the process of road traffic accident decision making. Once the relevant knowledge is acquired then it is represented in human understandable manner to arrive at the expected goals. Therefore, the knowledge based system was implemented by articulating the rules.

CHAPTER FOUR

KNOWLEDGE BASED EXPERT SYSTEM DESIGN AND IMPLEMENTATION

4.1. KBES Design

The second objectives this study is to perform best techniques, tools and approaches for expert system development. In chapter three (3) above best techniques (interview of domain expert and data mining techniques) was usually introduced and performed. This chapter involves the actual development of workable KBE system for road traffic accident decision making. Therefore, having all the necessary accident cases and the knowledge from automatic knowledge acquisition using data mining and domain knowledge from experts, the next task is coding the knowledge into computer using appropriate and efficient knowledge representation methods. For this research, web development tools such as PHP to compile the system on web, apache, and other supportive tools such as JavaScript and CSS frame work is used to develop the prototype. Relational database tools such as MySQL that serves as knowledge base (KB) is used to provide sets rule.

The following points shows reason why researcher chosen web development tools PHP, MYSQL as a knowledge base and C scripting language for the manipulation of acquired and extracted rules.

- All web development tools are open source and any platform supportive tools
- Simple to configure in multi way of using (if needed to available for online use and access) as an advisory system to determine future solution based on current accident.
- Efficient and reliable user interface design rather than other intelligent and expert tools such as PROLOG.
- Very simple to understand and access road traffic accident advisory service for user rather than logic tools such as CLIPS and shell.
- With web development tool since webpage used as a compiler not difficult for level of skill and knowledge gap if user can read and understand.

- It is simple to design knowledge base by using either relational database or built in that accurately represents relationships and sets of rule (collection of knowledge) respectively.
- Web based tools are best suited to solving complex arithmetical computations that logic program such as prolog could not did.
- Web development tools PHP is best suited to the current PC architecture (sequential execution) and are best optimized on parallel architectures that prolog programs are not.

The most important criterion for any KBES is the accuracy of their inference engines. When some information is missing, redundant and inconsistent it mentions the problem behind and explores solution. Web development tools could reduce such kind of risk rather than other logic program. Human readability helps the domain experts to understand knowledge content and even to modify its content. The system‘s dynamic nature to adopt a new knowledge from the user response is critically important to measure system performance. The knowledge based system implementation depends on the decision tree structure. Because of their simplicity, decision trees or generic rule-based models are commonly used to describe the knowledge acquired from domain expert. The if-then rules of KBS are generated based on the decision tree structure.

4.2. Building RTACKBES

The third objective of this study is to develop knowledge based expert system that make decision on road traffic accident problems and have good performance of usage. The goals of the systems are to insure that the final result of the system was usable and practically applicable for the end user. To achieve this goal, the KBS was divided into different sub system.

To achieve this objective the researcher constructed KBES (KBS) components for RTACKBES such as knowledge base, inference engine learning, explanation module and user interface shown in figure 4.1 below. The architecture of the RTACKBES shown in figure 4.1 depicts how the prototype works during road traffic accident decision making. As the new query (problem) is entered, the prototype of the system matches the new case to the solved case in the case base of the system by using rule based representation. If relevant cases are found within the database, then the prototype display the problems. Next, the prototype proposes a solution. The proposed

solution can be derived directly from a retrieved case that matches exactly or partially to the problem of the new case. Finally, the revised solution is retained in the knowledge base for future problem solving.

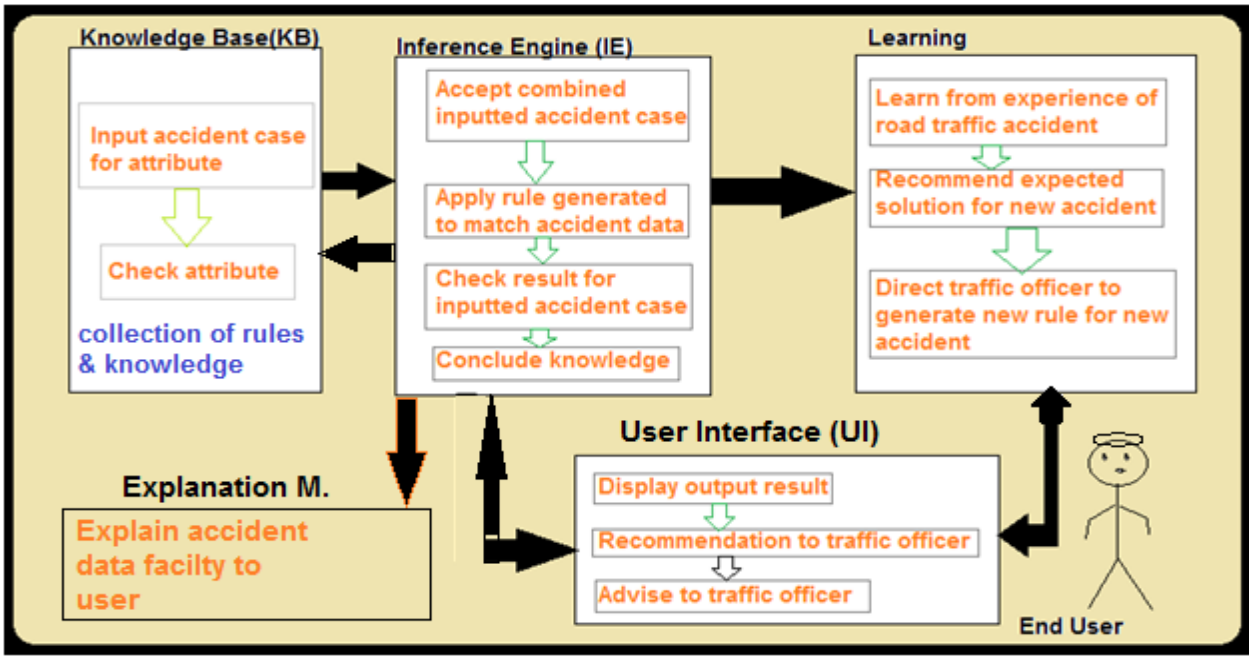

Figure 4.1 RTACKBES structure

4.2.1. The Knowledge Base

In this study the important part of knowledge based expert system was knowledge acquired from expert and accident data recorded on paper, which is structured to support decision making. The collection of accident facts, rules representation, and computational procedures that represent the knowledge based expert system are concerned.

As depicted in figure 4.1 the knowledge base incorporates the relevant cases that were acquired from the domain experts and record of accident data. The knowledge base of the prototype contains the domain knowledge which is used to identify the types of problems, symptoms and solutions in Road traffic accident management and control. Based on this, the knowledge base was constructed by using web development tools Apache, PHP implementation tool in the form of ‘if then rule’. The reason is that, web development tools are open source software and it is the preferred programming language for developing expert system. The advantage of expert systems over conventional programs is that their core algorithm is not encapsulated in the programming

code but stored as knowledge in an independent database called knowledge base and allows developer to develop an algorithm on relational database such as MySQL.

The acquired knowledge's are saved in plaintext file format. Case attributes are Accident moment, type of accident, road surface, road orientation, road surf, congestion, driver health condition, driver driving potential, driver driving on side, weather condition, pedestrian health condition, pedestrian walking on side, injuries of severity and preventive measurement are the important attributes chosen from all attributes. These attributes have significant impact on road traffic accident management and control process.

Ten rules are possibly generated by an expert for making decision. Thus, 44 rules extracted in the form of knowledge from data mining application and other important sets of rule represented knowledge from traffic experts are based on the number of preventive measurement that will be taken for future accident control. The following figure 4.2 shows that knowledge base learn from the facts given and using one of rule assigned for preventive measurement that if there are no signals, no traffic police on area where accident is happened " install enforcement and warning sign to limit driver speed". If facts are inputted and there are no rule organized so the system learning component produce message for user to save this accident case as new rule.
Finally, the number of facts depends on the number of rules incorporated into the knowledge base. Functionally, the facts in the fact base are used to compare against if (condition) part of rules stored in the knowledge base. In other words, if the given facts satisfy all the conditions which proved to be true, then the inference engine draw a conclusion. This is based on the pattern matching between the facts in fact base and their respective rules in the knowledge base.

4.2.2. The Inference Engine

Inference engine is the set of procedures for manipulating the information in the knowledge base to reach conclusions. The objective of the inference engine is to find one or more conclusions for a sub-goal or for a main goal of the consultation. It searches the facts and rules in the knowledge base then identifies and stores conclusions to use in new facts for subsequent inference.

From the two inference engine model forward chaining was selected because a multiple solution is expected from multiple accident causes inputted once into a knowledge base. In the propose

method the goal is to obtain the possible solutions that can be applied to different road traffic conditions. Because the goal (possible solutions) is unknown before the knowledge structure is developed and the input information is available, forward chaining can create an efficient structure of the knowledge that is understandable easily followed by the user.

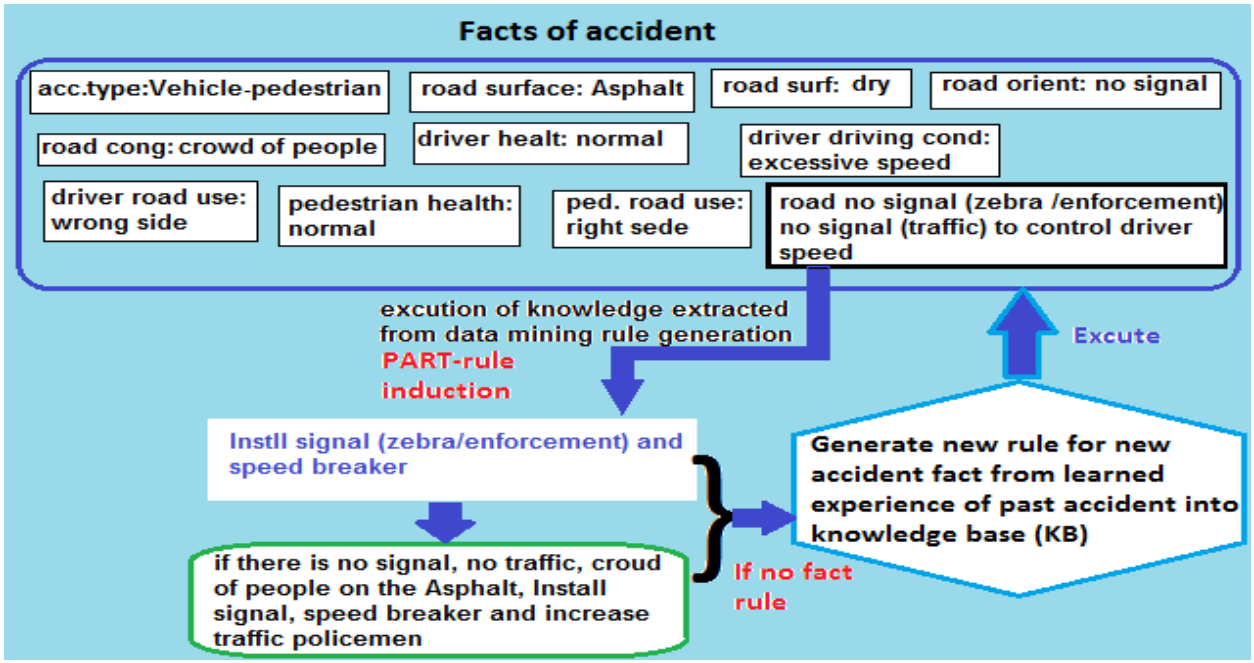

Figure 4.2: knowledge processing into and from Knowledge base and inference engine

For this study road traffic accident cases are inputted in to web based formed by the researcher. This form was used as a user interface for the road traffic police officer. Therefore, inputted road traffic accident cases processed by the rule 'if then' developed in relational data base MySQL writing C language by PHP for the similarity or dissimilarity of knowledge which is extracted from data mining results with inputted in knowledge base on web page.

4.2.3. User Interface

The user or the end user is the person for whom the expert system is developed and uses the system for its problem solving assistance. The domain expert may add new knowledge to the existing system while the knowledge engineer refines the knowledge in the system. The knowledge engineer converts a domain expert's specialized knowledge into sets of IF-AND-THEN-ELSE rules using instructions that a computer understands. The domain expert may be

the knowledge engineer and this may be of an advantage in the accuracy of the expert system. The user may not be a traffic engineer but it is recommended that the user be a traffic officer. The user seeks advice the system recommended after making decision. The researcher designed user interface that advises traffic officer to do task needed after and before accident happened. The designed RTACKBE system user interface with advising and guiding instruction on window is illustrated in figure 4.4 below. The welcoming window involve option that a traffic officer need to perform based on the accident happened. If the accident happened is crash of vehicle to vehicle traffic officer choose the icon "vehicle to vehicle".

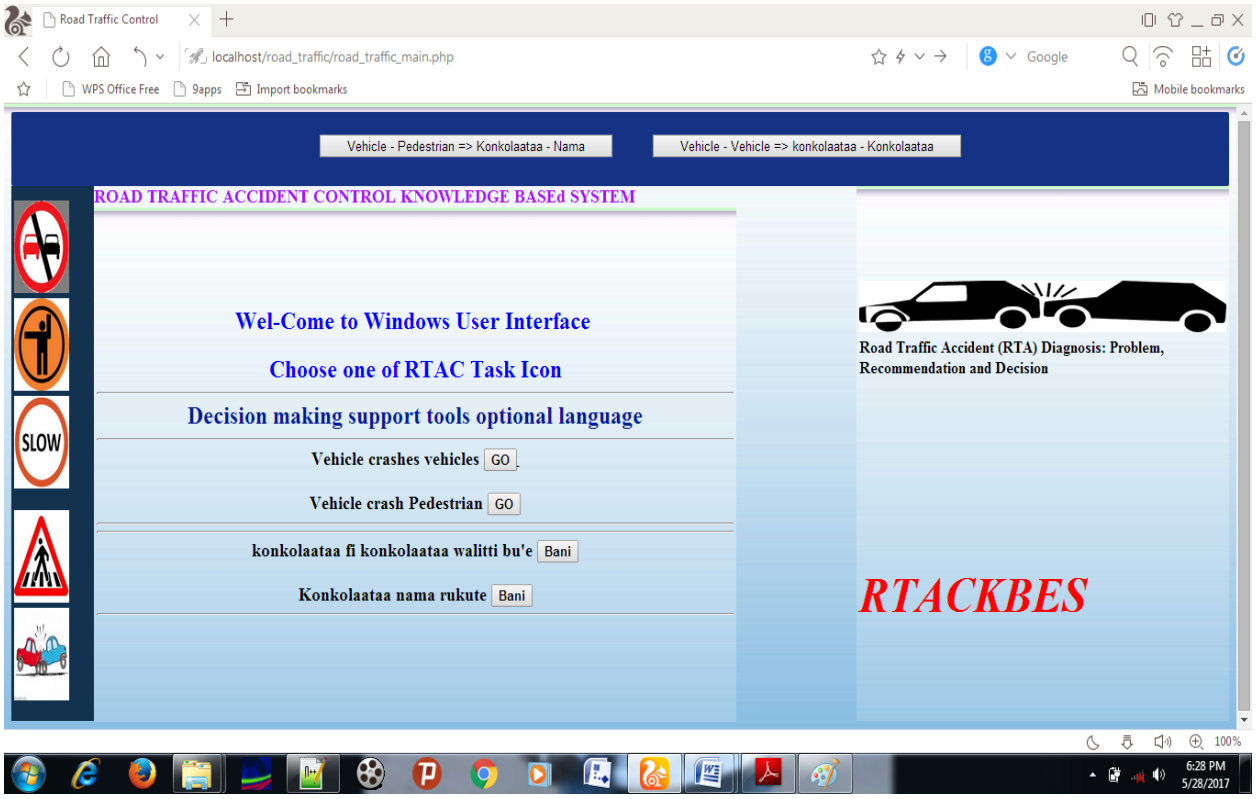

Figure 4.3 Welcoming Window of RTA KBES User interface

As depicted on the above figure 4.3 road traffic accident knowledge based expert system is specified clearly on web page using PHP and other web development tools like HTML, CSS, Java Script and Adobe.

The page was opened by using URL (universal Resource Locators) guided by hyper text transfer protocol (HTTP) by opening any kind of browser. Since the system is displayed on the browser

and supported by server the server name or static IP address will be given as HTTP://Localhost/ or HTTP://198.162.1.1/ . With opening this well come page the user can interact with the option of road traffic accident types task. This road traffic types are designed in a way user can click on simply and get the next task. The user identify the type of accident needs to be analyzed and if the accident happened was vehicle to vehicle crash the user choose "vehicle – vehicle or konkolaataa - konkolaataa" type of accident. If the accident happened was vehicle crash the person (pedestrian) the user choose "vehicle – pedestrian or konkolaataa nama rukute" type of accident.

As it is described in chapter two the researcher involved two optional languages in user interface that are English and Afan Oromo as shown in figure 4.4. The reason researcher used Afan Oromo was Jimma traffic police officer are more educated and familiar with Afan Oromo specially to analyze road traffic accident data. The language of current work and preparation of report was with Afan Oromo language. This task was processed with simple way of HTML link on the webpage using PHP code. As a result the user chooses one of road traffic accident type the next page is user inputting case in form. The user used the knowledge acquired and represented by PHP rule are illustrated in below figure 4.5

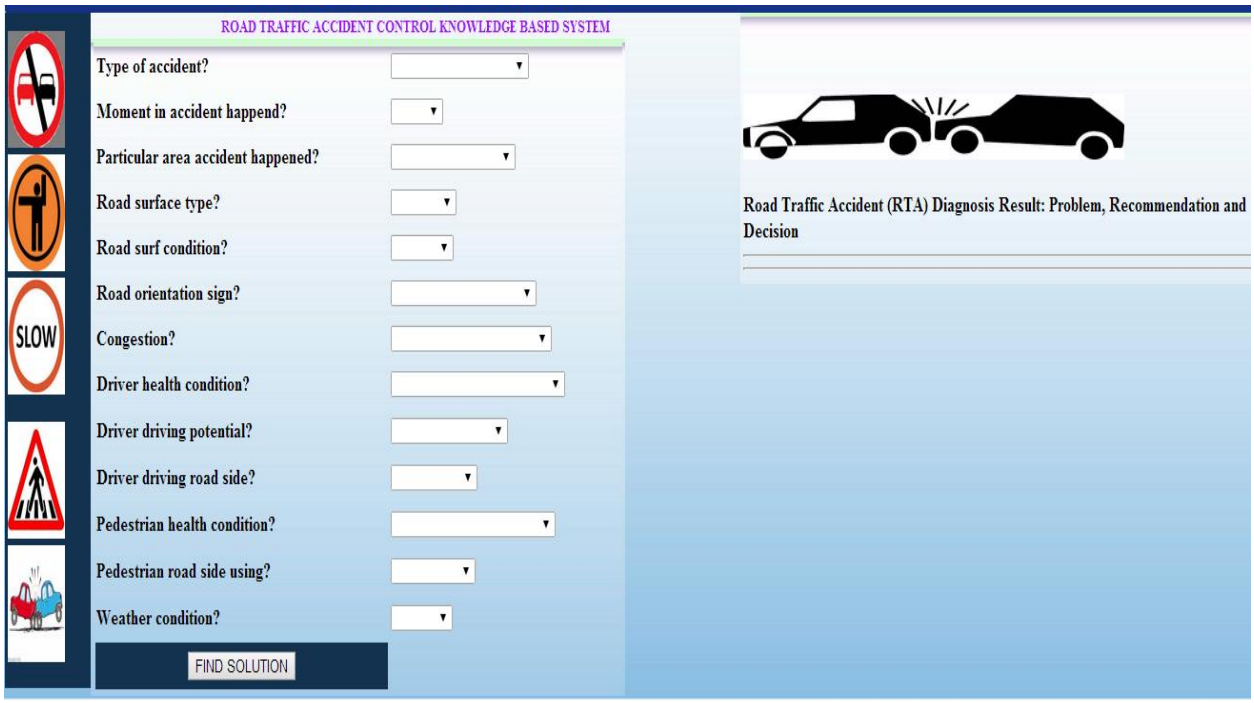

Figure 4.4 Window for case process into knowledge base for Vehicle crash pedestrian solution

In the above figure 4.5 the users interact with the window to analyze accident data after all about accident cases are collected by traffic police on duty. Figure 4.5 helps the user processing when vehicle crash pedestrian on the road. The window have different phase the user follow with instruction given by both English and Afan Oromo language. The option of language is important for user to process the accident data management system in efficient and effective ways because the traffic officers working in Jimma town are familiar with Afan Oromo language.

The first phase involves types of road traffic accident either accident is vehicle crash with vehicle or vehicle crash pedestrian who walking, crossing or running on road.

The second phase includes information about the road to identify weather the roads have been a risk factor that influenced an accident to happen. These road cases are road surface (road is asphalt or ground), road orientation (road have any traffic sign or not), road surf (road is pavement of wet, mud and dry) and the congestion (road is crowded with the people or car). If the road have problem or have no problem the user focus on the driver.

The third phase involves the driver situation including driver health condition that driver have physical or psychological problem, driver fatigue because of driving long period of time; driver driving potential that driver is driving with excessive speed, with no skill or without license; and driver driving road angle that wrong side angle, the right side angle and when crossing the road.

The fourth phase of the window case inputted for crash of vehicle to pedestrian involve the weather condition that can affect the road to have risk with driving on and affect driver to drive with inadequate visibility because of environmental factors (making it hard to detect vehicles and other road users).

The Fifth phase of case inputted window show the pedestrian condition. As researcher interviewed the domain expert sometimes pedestrian have physical or psychological problem are affected by vehicle accident. For example pedestrian who have eyesight problem, hearing problem are affected when walking on the road. Road users are affected with accident when crossing the road and walking without their right road side angle.

The sixth phase of case inputted window is about rank of injuries severity includes whether the injuries are high severe, medium sever and less severe. This rank of severity is help the traffic

police officer to make decision on the road traffic accident happened. If the injury severity is high severe the system generates the urgent solution for traffic police officer.

Finally the problem is identified and the system learns from new inputted case to make decision if there is the same case in the data base. If there is no such inputted cases the system store the new case with most future solution. The expected output is materialize in a systematic manner, the establishment of a database expert system for traffic accident problems would be appropriate to have an accurate analysis of the potential alternatives to be recommended. The vehicle to vehicle crash window shown in figure 4.6 below

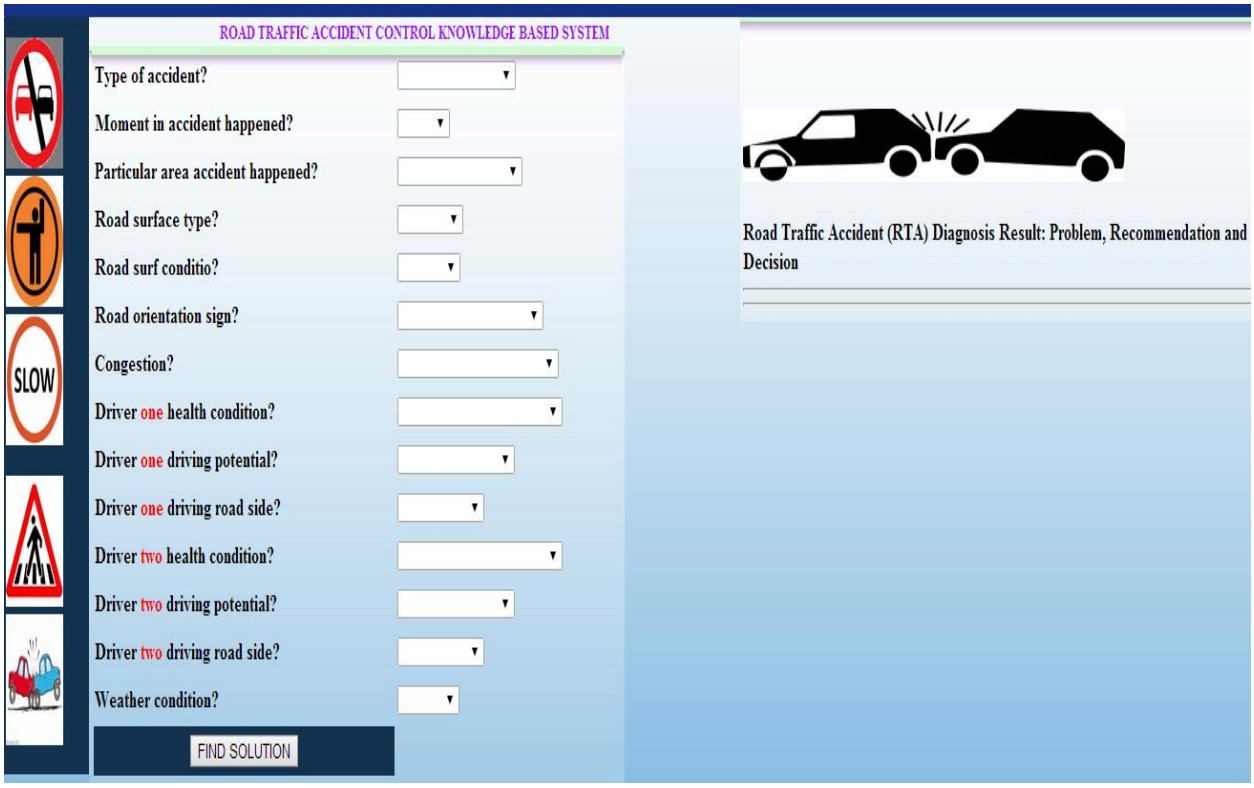

Figure 4.5 Window for case process into knowledge base for Vehicle to Vehicle crash solution

The difference between figure 4.5 and figure 4.6 is the absence of pedestrian information since the user process the vehicle crash vehicle road traffic accident.

In addition figure 4.5 and figure 4.6 are different in concern of the third phase which involves the driver situation include driver one and two health condition that driver have physical or psychological problem, driver one and two fatigue because of driving long period of time; driver one and two driving potential that driver is driving with excessive speed, with no skill or without license; and driver one and two driving road angle that wrong side angle, the right side angle and

when crossing the road. In figure 4.6 the driver that can be cause for accident measure in both angle that driver one and driver two as illustrated in figure 4.6. After the case is selected by the user based on the accident happened in both figure 4.5 and 4.6, the user find solution for the accident. The solution includes recommendation, preventive measurement and future solution for probability of accident.

4.2.4. Learning component

As it is discussed in section 4.2.1 and 4.2.2 the system advises the user learning from the inputted traffic cases. The system has two learning component that are advisee and guiding components. If a rule is stored in the knowledge base, the system advice the traffic officer with the solution. The second component is if the knowledge base not has the knowledge about the case inputted, the system guides the traffic officer to save that new case. This means the system is not case based, but the rule structured in knowledge base was 44 that are not enough for some unique accident happening probability.

In addition if there is no similar knowledge representation rule in the knowledge base the system recommends the user that the case is new which the system not involved in knowledge base. The system learns from the experience by requesting and supplying optional solution for new accident inputted in knowledge base. The learning components of the RTACKBES is illustrate in figure 4.6 below. The conversation between traffic officer and system when the accidents were new;

- The user inputted the cases and tell system “find solution”
- System search solution through inference engine (IE) manipulation task and respond “there is no future solution generated but learned from past experience”
- Then the learned result can be: “display suspected solution for new inputted accident cases”
- System ask the user “if you would agree to this solution you can generate complete rule for new knowledge” by choosing “**Yes**” or “**No**” option. “**Yes**”: represent severity is high so store new knowledge into knowledge base, “**No**” severity is less so not required to store.

For example if the user inputted new accident case for vehicle crash pedestrian accident type the following figure 4.6 and 4.7 depicted machine learning component interface:

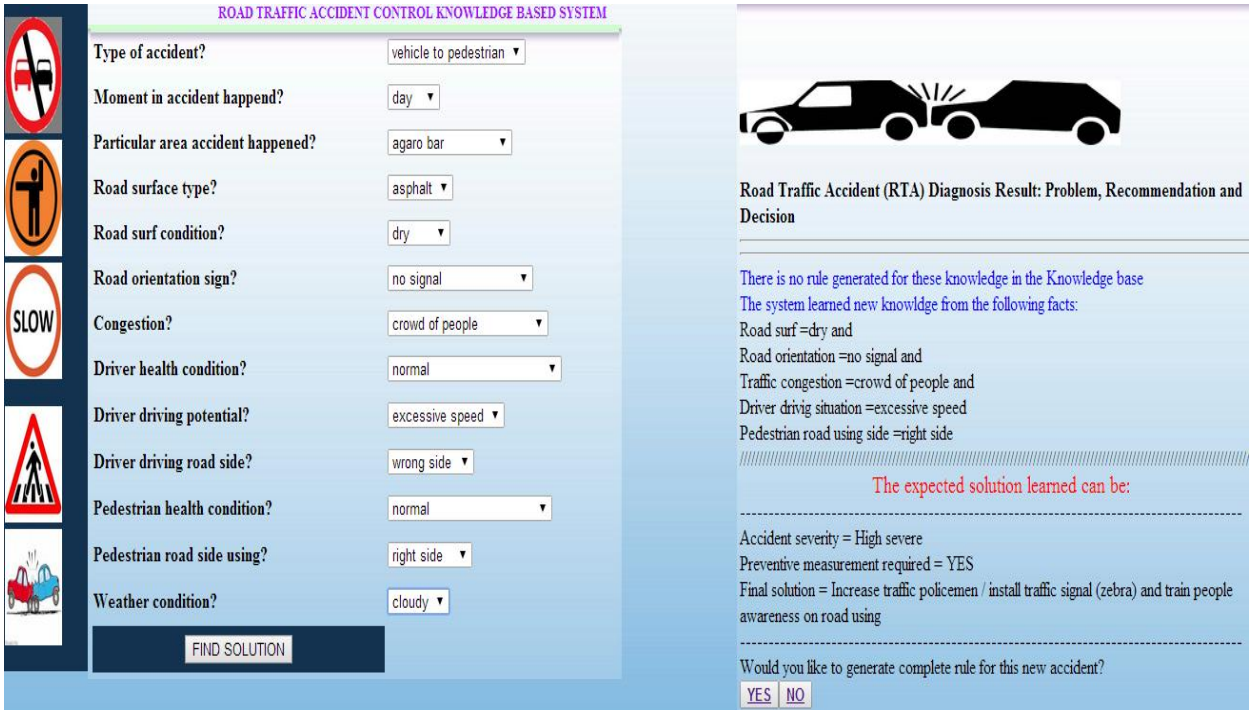

Figure 4.6 Learning components for new case and system advice

It also request the user to generate the new accident case that represented sets of rule extracted from the experience of traffic expert. The new case saved with the following window figure 4.7.

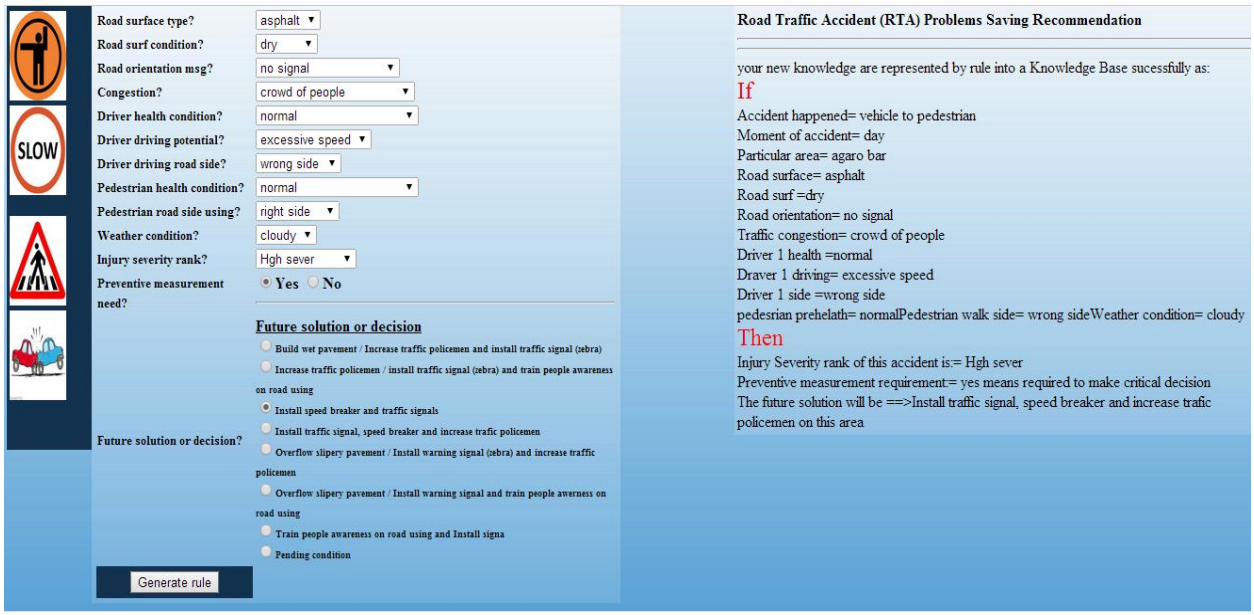

Figure 4.7 New accident case representation windows

The most remember able in this section is the case just serves as a new knowledge represented by a set of rule in the knowledge base. This is considering that what traffic officer can do if there is no knowledge represented with the rule representation by either extracted from data mining or from traffic expert in the knowledge base. In the above figure 4.7 after the new case are inputted into user interface the traffic officer request system to generate rule representation in knowledge base. The system responds that new accidents are stored successfully into knowledge base and display the rule representation of knowledge from knowledge base on user interface.

4.2.5. The Explanation Module

The Explanation Module contains explanations for every inference made or piece of advice given. All procedure in which the road traffic knowledge base processed is also explanation module. Foremost one of the interesting features of knowledge based systems is the ability to explain itself. The knowledge based system developer used the explanation module to have more realistic dialogs with the system user. The 'how' feature of the prototype is incorporated so as to give explanations how the system reach to certain conclusion.

Defining knowledge structure in web based tools are done by using case input for vehicle to vehicle crash and vehicle to person crash windows. It is very easy to define knowledge structure with PHP. Because it is simple to add attributes in description of case structure and set properties of attributes or metadata of attributes. Metadata of attributes are weight of attribute, data type of attribute and similarity function. During configuration of knowledge structures, PHP creates codes by developer and saved in PHP file format. Most significant attributes are set by declaring higher weight as compared to other weights.

As the knowledge is extracted from data mining application or traffic expert final solution has discovered as decision in the knowledge base. Example: if road surface = asphalt AND road surf = dry AND etc. THEN one of the following final conclusions advised and recommended by system are:

- ✓ Build wet pavement at road side and Install signal(direction)
- ✓ Install barrier for pedestrian (zebra) and driver limit speed
- ✓ Install barrier (Zebra) for pedestrian and no pedestrian cross road
- ✓ Build warning sign (Stop / direction) and Increase traffic policemen

- ✓ Install warning sign (Stop / direction signal)
- ✓ Overflow slippery pavement remove drainage and build warning sign
- ✓ Train community how to use road
- ✓ Increase number of traffic policemen
- ✓ Pending the solution where the need of preventive measurement is 'no' and severity rank is "less severe".

4.3. Building RTACKBES prototype

As the fourth objective of this study that is to develop a prototype expert system for road traffic accident the researcher built prototype from acquired knowledge, using web development and data mining techniques. As mentioned in chapter 3 of this study web development include three basic common tools at once that Apache, PHP and MYSQL database. MYSQL database serve as knowledge base. Extracted rule from data mining PART rule induction algorithm was written with C scripting language using Dreamweaver. Web development tool like PHP used to write C code for the manipulation of set of rule in the knowledge base by configuring it. Once knowledge based system is constructed on web based tool it was corporate once user could interacted to system for rod traffic accident advisory service. This could by inputting an acquired accident facts, retrieve manipulated fact result and recommendation.

In both figure 4.6 and 4.7 as the user inputted the cases in all field related to the database attribute inference check whether the inputted value in knowledge base and value in database is similar. The final solution processed based on the problem, preventive measurement and decision as a solution.

4.3.1. Managing Methods

The method library stores classes that actually resolve the task. These classes can resolve the knowledge based cycle using in programming or using GUI. All tasks that are mentioned above should have their own methods to be assigned in order to achieve the tasks goal. The following are lists of methods which are used to solve tasks for this KBES for RTAC application.

Load knowledge Base Method: This method returns the whole available knowledge from the knowledge base to designer. In this study the knowledge base is loaded through rule developed

in MySQL as a knowledge base using web page. Road traffic accident characteristics that show attribute values also load.

To summarize the chapter, the acquired knowledge from the domain expert was represented using a set of rules but all the problems are not included in the knowledge base because the researcher is required to design the prototype. The rules were constructed in the prototype with if-then format. If- then format of the prototype mainly depending on the decision tree model, which was discussed in chapter three. To achieve the goals of system, the inference engine using backward chaining reasoning mechanism. Furthermore, the knowledge based systems have different components as indicated in the knowledge based system architecture. The knowledge based system components are knowledge base, inference engine, learning, user interface, and explanation module.

CHAPTER FIVE

5. PERFORMANCE EVALUATION AND TESTING PROTOTYPE

As discussed in chapter one the main objective of this study is developing knowledge based expert system for road traffic accident data analyst and decision maker. After implementing knowledge based expert system the final step is measuring the performance of the system to know the extent to which it has achieved the objective or not. For the purpose of this study, road traffic accident control KBES is tested and evaluated based on the objective of the system. This is to measure the accuracy of the system during the problem solving processes. The measurement is from the point whether the system achieved its objective or not. In this study the performance of the system was measured against human expert decision using predictive validation method and user acceptance of the system was carried out during system user interaction.

5.1. Testing RTACKBES and evaluating the performance

The fourth and last objective of this study was to evaluate strengthen of RTACKBE system beyond solving current road traffic accident problem, predict forward research and conclusion. To achieve this objective the developed expert system investigated the road traffic accidents management and control scheme that was not studied in Jimma town at present. This covered all traffic congestion problems and traffic accident incidents that took place. Traffic control measures will now be analyzed by using rules for inference by forward chaining to achieve potential alternatives. This is more preferable system than the current manual system.

The KBES user acceptance was measured by using open and close ended questions. The system evaluators directly interact with the system to measure its performance from the points of its correctness in providing solutions for different road traffic accident problems. In addition, the validation test was done by comparing solved cases against the system conclusions on the similar issues. By comparing the result obtained from the system conclusion, the evaluators determine the performance of the system. Therefore, knowledge based expert system performance is measured first by using open and close ended question, and second by using the test case

validation method. The details of each evaluation method were discussed in the following section.

System user interface explanations of every icon are designed with user understandable languages that are Afan Oromo and English. The system functionality was as just as the need of the traffic police officer. The accessibility of system is efficient and flexible.

5.2. User Acceptance Evaluation

This is a part that evaluates RTACKBES in how much a user accepts it comparing the quality of previous manual road traffic accident data analysis system with new developed system. The importance of KBES developed for road traffic accident decision making by using web page development tools that Apache, PHP and MYSQL was evaluated. The researcher determined to make the way a user evaluated either accepted or not the system in standard ways.

User acceptance was evaluated in terms of similarity on system case accepting, case processing, using and reusing the case of knowledge base between developed KBES and previous manual system.

As the road traffic accident data is collected by traffic expert from the accident happened traffic officer was evaluated the system concerning the following tasks.

- Open the welcome window of KBES its content,
- Open case inputting knowledge base window and input all collected accident cases which is exactly the same with hard copy format,
- Make decision that includes problem, recommendation and decision automatically on solution screen. The solution is exactly the same with a solution that is made by traffic police officers during the accident happened by using manual.
- Using and reusing this case for another accident.
- Close and open with the same steps

To ensure this knowledge based expert system visual interaction evaluation method allows the domain expert to directly interact with the system. Direct interaction intend system user to evaluate the performance of the knowledge based expert system from the users point of view. In

addition, this method helps to ensure the performance of the prototype by assessing the feedback acquired from the domain expert towards the developed system. A questionnaire was developed to assess and evaluate the appropriateness and applicability of RTAC KBES in the domain area.

The system evaluator was traffic officer in a way researcher explain the use of system. In order to achieve user acceptance evaluation measurement experienced traffic police in addition to traffic officer selected as evaluators. Thus 4 traffic police officers and other 4 traffic police men experienced in traffic accident decision making process were participated in the evaluation. So the total number evaluator was eight (8) traffic police.

The user acceptance evaluation was carried out by using questionnaires. The questionnaire includes both close and open ended questions by Afan Oromo language. Using these questions evaluators directly interact with knowledge based system after researcher gave orientation how to use the system. Totally nine (9) evaluation questions were prepared for system evaluators. The questions are divided into two parts, the first five (5) questions are close ended questions which helps system evaluators to check on the user interface design aspects, easiness of the system to use, attractiveness, correctness of the decision, adequacy of knowledge content, the problem solving ability and significance of knowledge based expert system in troubleshooting road traffic accident problem diagnosing services. On the other hand, the remaining four (4) questions are open ended questions which used to collect constructive feedback from the system evaluator‘s based on the system‘s conclusion.

Among traffic police officer the evaluators were allowed to rated the options as excellent, very good, good, fair, and poor for these closed ended questions. Therefore, for easiness of analyzing the relative performance of the prototype based on the user evaluation after the interaction with the system, the researcher assigned numeric value for each of the options given in words. The values are given as Excellent = 5, Very good = 4, Good = 3, Fair = 2, and Poor = 1. Based on the given scale, system evaluators provide a value for each close ended question. Thus, this method helps the researcher to manually examine the user acceptance based on evaluator‘s response. The user acceptance of the system is measured manually as follows (Aboneh. 2007):

$$\mathbf{AVP} = \sum_{k=1}^{n} \mathbf{SVi}\ \frac{\mathbf{NRi}}{\mathbf{TNR}}$$

Where, SV scale value, TNR total number of respondent and NR is number of respondent. To get the result of user acceptance average performance is calculated out 100%.

$$(\mathbf{AVP} = \sum_{k=1}^{n} \mathbf{svi}.\frac{\mathbf{NRi}}{\mathbf{TRN}}) \mathbf{x} \frac{\mathbf{100}}{\mathbf{NS}} \%$$

Where NS is number of scale and AVP is average performance. The following table summarizes the results obtained from the respondents'. The question is translated to Afan Oromo as shown in Appendix II.

Table 5.1 User Acceptance Testing

Question for Evaluation	**Performance Value**						
	1	**2**	**3**	**4**	**5**	**Average (AVP)**	**AVP * 100%**
How was Adequacy and clarity of result for decision making?			1	4	3	4.25	85
How was Relevancy of the retrieved cases in the decision making?				4	4	4.5	90
How exactly Fitness of the final solution to the problem at hand?			1	2	5	4.5	90
How much ease of use of the RTA KBE system?				1	7	4.875	97.5
Is that attributes relevant with Road traffic cases represented?			1	4	3	4.25	85
How exactly efficiency of the system in time?			2	2	4	4.25	85
Is the resource was adequate to the system?		2	4	2	0	3	60
How were interactivity of the user interface?			2	3	3	3.875	77.5
How you rate the significance of the system in the domain area?			1	3	4	4.375	87.5
TOTAL AVERAGE						**4.21**	**84.17**

As shown in table 5.1 above, the respondents' rated for RTACKBES as good to the question "How was Adequacy and clarity of result for decision making?" was 12.5%, Very Good 50% and 37.5% was excellent. Similarly, for the question how was relevance of the retrieved cases in to support user's decision making rated very good by 50% of the respondents whereas the remaining 50% of the respondents rate it as excellent. In the case of fitness of the final retrieved

solution to the new problem at hand around 25% of the respondents rate the prototype is very good whereas only 62.5% of the respondent's rate as excellent and other 12.4% was good. All the respondents' rate ease of use of the system is 12.5% very good and other 87.5 were excellent. 12.5% of the respondents rate the relevance of attributes in representing accident case rate were good, 50%very good and the remaining 37.5% of the respondents' rate as excellent. Lower rating is assigned to the prototype efficiency in terms of time and the resource it has. Only 50% of the respondents' rate the system as excellent in terms efficiency in time as well 25% good and very good. Around 25% of the respondents' rates the available resources are fair and 50% good with promise and 25% of the respondents' rate the resources available are very good. The user interface of the prototype is also rated by respondents it is good around 25% and the remaining 37.5% is assigned very good and 37.5% excellent. To this end 50% of the respondents rate the applicability of the prototype in their domain excellent and the remaining 12.5% and 37.5% of the respondents good and very good respectively. Finally, based on the evaluation of all the respondents the average performance of the prototype is 84.17%, which is above very good. This performance result shows the prototype has a promising applicability for Road traffic accident decision making process.

In general, the significance of KBES in road traffic accident control process has achieved an encouraging result (result expected to achieve). In addition, relevance and clarity of retrieved cases also rated by the respondents with a highest rating value. One interesting feedback provided by the domain experts is adding explanation facilities of the system. According to the domain experts, explanation facility is important in the adaptation of the retrieved solution. If users of the system get a more explanation about the retrieved solution and the problems itself, user can easily decide whether to use or not use the retrieved solution.

Regarding the question of knowledge coverage issues by RTACKBES, the respondents replied that the knowledge based system is efficient to respond from the stored knowledge in the knowledge base.

Similarly, the respondents were asked to express their opinion for the question of "Does developed knowledge based systems replace the tasks of road traffic accident problems diagnosis with relevant solution?" Most of the respondents confirmed positive feedback. As their response

indicated, as far as the knowledge base contain all the relevant knowledge acquired from the domain expert, the knowledge based system provide similar decision as human expert. Even the system can be considered more flexible due to its reusability of the existing knowledge in a better way than the human experts. The truth is that, human expert consider different conditions during the road traffic accident control processes involving road, driver and pedestrian factors.

At the same time, most of the respondents replied that the knowledge based system contains adequate knowledge for all rules incorporated in the knowledge base. As their suggestions indicated, the system must update the existing knowledge and learn from the environment. To handle such issues the researcher incorporated single case to dynamically update the existing fact in working memory. This component helps knowledge engineers to design a knowledge based system that update the fact base dynamically. But, it is clear that incorporating learning component for rule based system is difficult and complex as the numbers of rule expanded.

Therefore, for specific domain area incorporating such component will improve the efficiency of the system. For the question that concern about the significance of the system in the domain area, all the system evaluator‘s responded that the system add value in the domain area. Therefore, the application of new technology simplifies the working environment in the domain area. In addition the system also can reduce the burden of human expert by saving their time and energy spent while referring recorded accident and other manuals to solve the problems. Not only the system recommends solutions, but it can also contribute for the reusability and sharing of the same knowledge for different purposes such as road construction.

5.3. RTACKBES Validation by Using Test Cases

In Road traffic accident control decision making adaptation is a commonly required task. Since this research main goal is developing road traffic accident diagnosis and solution based decision making, adaptation is necessary. Four Traffic policemen who are different from four traffic officer (traffic expert) have experience and expected in road traffic accident problem solving was test case. The validation of system measured by the knowledge (information collected while accident) constructed in knowledge base by the rule and accident data at hand. Researcher prepares knowledge base knowledge which is more relevant with manual use of knowledge. The

difference and similarity of cases which traffic officer are used to make decision doing on paper and doing expert system was the major RTACKBES validation. Therefore, the adaptation has to be performed manually by a human traffic expert as shown in the figure 5.1 below.

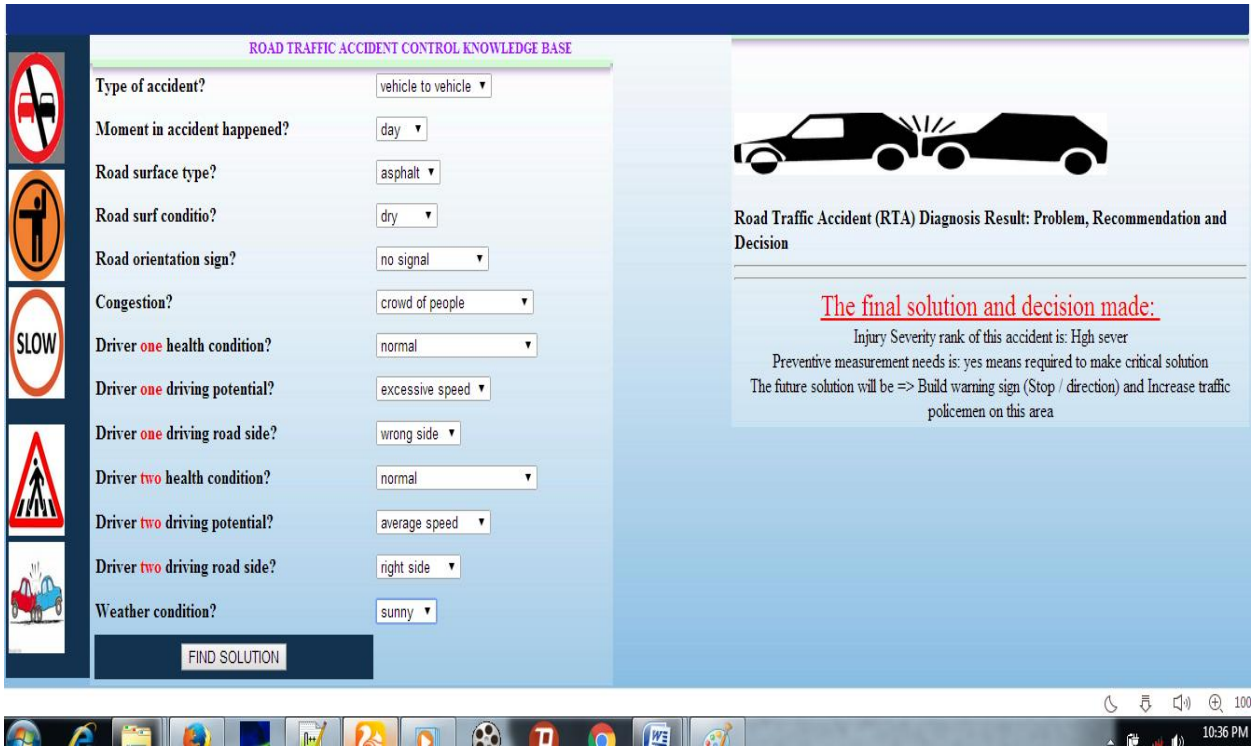

Figure 5.1 Revision and case check Interface for vehicle crash vehicle

The KBS testing procedure was carried out by system evaluator to evaluate the solutions suggested by the system were correct or incorrect. System evaluators compared the decisions made by the system against human expert. Then system evaluators validated the number of correct decisions made by the system. The result of the comparison shows that the rule based system has made close decision in the troubleshooting process of problems as human experts did.

Since save case categorized and based on the type of accident figure 5.2 show that window of revision and case check for vehicle crashes pedestrian.

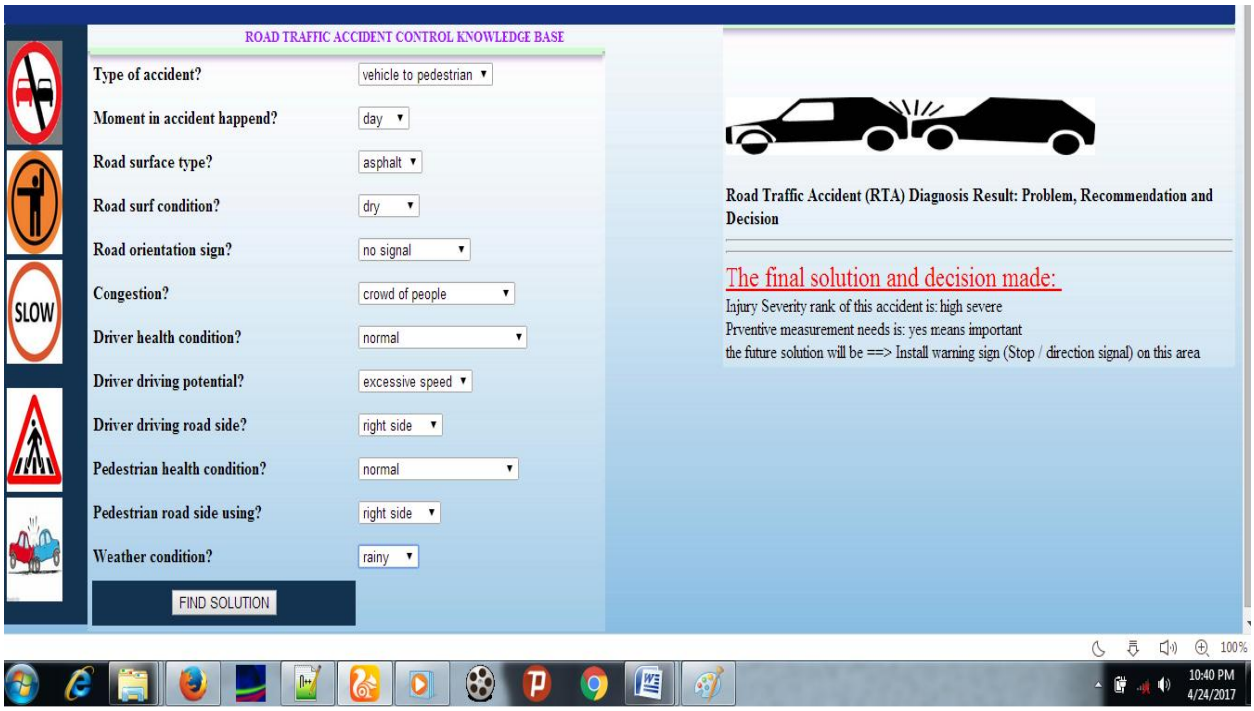

Figure 5.2 Revisions and case check Interface for vehicle crash pedestrian

5.4. Decision Variation between System and Human Expert

As the main goal of this study is to achieve on improved decision making process by developing RTACKBE system human expert knowledge processing and decision making tasks are converted to system expert.

As discussed in section 5.2 above the decision made by the system have shown difference with decision of human expert during test case validation. There are different contributing factors for the variation of decision made by the knowledge based system. First, the KBS is limited to the knowledge incorporated in the knowledge base. The main problem is the knowledge available for road traffic accident controlling. This mean sometimes the problem may be driver within vehicle crashes pedestrian accident problem but the expert system suggests other solutions such as, pedestrian using wrong walking side, traffic sign and enforcement and to environmental factor. For example, for the problem the drivers have no license; there are different cases. if driver is driving slowly and crash pedestrian without any sign on road that expose driver is driving with excessive speed, most of traffic officer go to focusing on pedestrian using road, Environmental situation and road have no sign or enforcement for the pedestrian cross road. Due

to inability to understand the context of the different cases the system suggests different solutions this may complicate the way of solving problems. Therefore, if the problem and cases are not clearly specified, it will expose and lead to waste time and resource improperly. As a result, the case and solution can be considered differently by different domain experts. Such issues affect the accuracy of the system during the test case validation.

5.5. Discussion

The respondent was responded on the importance of the system by evaluating simplicity of RTACKBE system in terms of its efficiency and ease of using is 97.5%. This result involves functions that traffic officer can access and use with optional language (English and Afan oromo) in efficient and effective way.

As a result of Africa (2011) developed expert system algorithm for Computer System Diagnostics using **PHP** evaluated system performance was 93.3%. Even though this study performance was 84.2% opposite to this study Africa (2011) used interview technique to acquire knowledge for his system prototype design. In this study the researcher used both interview and data mining techniques for knowledge acquisition, modeling (J48 classification algorithm) and representation (PART rule induction classification algorithm to determine sets of rule).

The prototype knowledge based system is promising and applicable in the domain area with the user acceptance of 84.17%. The feedback and suggestion of domain expert revealed that the knowledge based system satisfactorily gained user acceptance. The system acceptance evaluations used open and close ended questions to directly interact with the system.

The user accepted RTACKBE system as a result of user acceptance evaluation and performance results are achieved the objectives of the study. Performance evaluation question was distributed to the human expert police officers and acceptance of the user focus on the similarity between the cases collected and knowledge base. The system acceptance evaluations used open and close ended questions to directly interact with system. The following 5.2 table shows system evaluators responded result help to achieve on the objective of this study.

Table 5.2 summary of system evaluators Result

Domain expert result	**Total Respond Result**	**Average out 100%**
Poor	0	0
Fair	2	5
Good	7	17.5
Very Good	25	62.5
Excellent	33	82.5
Total average performance	**421**	**84.2**

As it is illustrated in the table 5.2 above, based on system performance gained from user‘s visual interaction with the prototype Knowledge based expert system using the closed ended questions; No evaluators respond as poor but as fair two times (5%), good seven times (17.5 %), very good twenty five times (62.5%) and excellent thirty three times (82.5%). The total average user acceptance evaluation result of knowledge based system was 84.2%. This result shows that more than 84% users accepted the correct decision, knowledge content, learning ability, contribution and significance of the knowledge based expert system. Therefore, the prototype knowledge based expert system satisfactorily makes the right decisions and provide appropriate recommendation for road traffic accident process.

As a result of Deguzma (2015) knowledge based expert system development on intersection improvement decision made by traffic expert. The components of the system were traffic accident and congestion. The performance was not identified but the research achievement is specified. Programming language used was web development tools such as PHP and MYSQL as a knowledge base (KB). The technique employed was interview and record of traffic accident document without application of data mining. Opposite to this study it was limited with traffic expert decision only when accident happened on intersection, so the solution is limited. In this study the decision support vehicle to vehicle crash and vehicle crash pedestrian everywhere.

For a cases of knowledge base RTACKBE system displaying solution improves RTACKBE system with evidence from human experts and data mining result. The response of open ended question and performance evaluation of developed RTACKBE system shows the exact

improvement of solution needed for road traffic accident problems traffic police officer need to analyze and make decision. The human expert deal comparison between human expert problem solving techniques and system expert problem solving techniques for road traffic accident problem. In concern of respondent exactly fitness of the final solution to the problem at hand and similarity are 90%.

Adding cases improve RTACKBE system with evidence from human experts and data mining result. Adding case into knowledge base made by researcher is critically to improve the RTACKBE system. In determining the cases of solution and preventive measurement it needed comparing human expert and system expert with the term that an attributes relevant with road traffic cases represented are similar with human expert of road traffic accident is 85%. The cases are defined in both language that English and Afan Oromo that support local police officer to evaluate problem and make decision.

The promising result is achieved in the study. The reasons that the prototype couldn‘t achieve 100% retrieval and reuse performance could be due to the data and the algorithm used to develop the prototype. Specially attributes quality of road traffic accident controlling and the technique currently used in decision making process. Writing the hard copy of data and making it more clearly in term of system also affects system prototype.

Similarly, the response of system evaluators confirm that developing such prototype knowledge based Expert system can reduce the existing knowledge gap in areas where road traffic officer were unable to make decision on road traffic accident problem or where there are probable of no expert. As discussed in section 5.2, the system validation was carried out using solved problems by human expert. As the result indicated, the prototype has made same decision with domain experts in the processes of troubleshooting the problems. The causes for the difference between human expert and the knowledge based system in decision making processes identified. The system also contributes valuable roles in the process of sharing experiences and skill in road traffic accident controlling mechanisms.

In general, the KBES approach in designing road traffic accident control and decision making system shows an encouraging result for retrieving relevant cases and proposing solution so as to

give decisions for new road traffic cases. It also attain promising user acceptance as it is evaluated by the domain experts. The domain experts (evaluators) assign more than average value for all parameters that are used in the user evaluation form for the prototype. This shows that the prototype achieves an encouraging result from domain expert side in retrieving a ranked order of relevant cases, as well as in proposing a solution to new accident cases. More over the prototype achieved promising result for its speed and easiness to use from the perspective of domain experts.

The performance of the prototype knowledge based system has got good user acceptance by the system evaluators. As a result, the knowledge based system would assist human expert to provide a better road traffic accident control service. Therefore, this system helps the general users to solve problems occurring after road traffic accident happened in their own instead of waiting for technical and accident information analyst.

The expected output of the study shown in chapter four of figure 4.1 materializes in a systematic manner. The establishment of a database expert system for traffic accident problems would be appropriate to have an accurate analysis of the potential alternatives to be recommended.

CHAPTER SIX

6. CONCLUSION AND FUTURE WORK

6.1. Conclusions

Nowadays developing system that contributes in road traffic accident management and control attracts many researchers from data mining and AI application. To address the RTA problem and improve solution on the area where road traffic accidents are increasing, data mining application researcher conducted research that predict the severity of injury from historical data stored in MySQL that act as a knowledge base. Therefore this data is processed by WEKA software due to selection, preprocess, transform and stored in to MySQL database (knowledge base). This applies only the prediction classification using WEKA model algorithm. This data mining application is important and strong in classification algorithm with the WEKA user interface without design of a user interface.

Data mining and AI application in corporation are powerful research field that attract researcher to develop system prototype help a user input cases into a knowledge base, system learn from inputted case and make decision using any knowledge representation such as rule, case and both rule and case representation.

In this study the researcher used both data mining application and KBES which is a collection of Artificial Intelligence (AI) techniques that enable a computer to assist people in analyzing specialized road traffic accident problems. Researcher used data miming application for data preparation, processing, preprocessing (data cleaning), noise removal and check validity of data using WEKA experiment and models. This was because clean data extraction is important. The study was conducted having the main goal of developing a prototype KBE system for road traffic accident decision making by using manual and automated knowledge acquisition techniques that can assist the domain experts.

All parameters of the problem that was required to build the system have been defined and designed using web development tools. The traffic rules for the KBE inference system were extracted from traffic experts and historical data. In order to test and evaluate the proposed system, it has been demonstrated for a case study of Jimma town traffic.

During the prototype development, real world road traffic cases are used from two station of road traffic office of Jimma town after passing through KDD Model using WEKA 3.6.5 software. For knowledge base web based tools such as PHP, MySQL and Apache software was used to develop prototype.

From the developed KBE system prototype traffic police officer can simply generate decision for road traffic accident happened from the knowledge stored in knowledge base. The database contain road traffic data imported from WEKA (file format (*.arff) converted to file format (*.csv)) to MySQL database. KBE system prototype linked to MySQL data set and retrieves it for user. Finally the user inputted cases in KBE system prototype, system learned from the inputted cases and make decision that a user wants to see.

When measuring the performance of the system, promising results are found. The standard measures of information retrieval are used to measure the retrieval performance of KBES. The total average performance of promising result to apply KBES for road traffic accident decision making process is 84.17%. This result also indicates the performance of the system evaluated by the potential users' of the system and achieved 84.2% performance.

While doing this study, the collected road traffic accident cases are not enough in size and do not have all independent attributes and features used by road traffic experts road traffic accident control and management process. In addition, the general knowledge explanation facility to advise the user when similar previous solved cases are not found in the case base is not achieved in this study.

6.2. Recommendation

The target of researcher was to develop road traffic accident decision making KBE system for improving decision making process and road traffic accident data analysis task. Therefore, different problems and challenges were suggested while using models, tools, representation, etc to be investigated in futures.

The developed RTACKBES was expected to perform the task of decision making process. As a result of finding obtained from the evaluation made by the experts on the prototype KBE and results generated by system the following recommendation to become forwarded.

- RTACKBE system is developed to improve decision made by traffic police officer and all cases involved in this system support decision made by traffic police expert. In addition decision making process on vehicle to vehicle crash and vehicle crash pedestrian accident, a system that involves the function for other stake holders such as road traffic engineer, road designer and for students will develop RTACKBES.
- The knowledge for development of RTACKBE system was represented as rule based. Rule based systems solve problems from scratch, while case based systems use pre-stored situations to deal with similar new instances. Therefore, the integration of rule based reasoning with case based reasoning would solve the limitation when representing knowledge in the form of if then rules unable to draw a conclusion.
- Inference step model is the most important procedure the researcher followed to develop KBE system. For this study forward chaining model of inference engine is applied for more than two or many in number possible goal. It is important to applying backward chaining model when the possible goals are known and reasonably small in number.
- Road traffic officer sometimes needs system result or output in hard copy in the form of report for head office. Design of RTACKBES user interface need to involve feature support format like print, report and authentication.

Reference

Abraham, A. (2005). Rule-based Expert Systems. In P. H. Sydenham, & R. Thorn (Eds.), Handbook of Measuring System Design (pp. 909-919)

A.D. M. Africa (2011); an Expert System Algorithm for Computer System Diagnostics: International Journal of Engineering (IJE), Vol. 5, No. 5, PP. 435 -467

A. Salem. (2007)Case Based Reasoning Technology for Medical Diagnosis.‖ World Academy of Science, Engineering and Technology

Beshah, T. (2005); Application of data mining technology to support RTA severity analysis at Addis Ababa traffic office: Addis Ababa, Addis Ababa University

Bianchini A. (2012); Fuzzy representation of pavement condition for efficient pavement Management: Computer-Aided Civil and Infrastructure Engineering

Chang, M., A. A., et al. (2005); "Traffic Accident Analysis Using Machine learning Paradigms" Informatica 29(1)

Cohn L. F. and Harris R. A. (1992); "Knowledge Based Expert Systems in Transportation", NCHRP Synthesis 183, Transportation Research Board, National Research Council, Washington D.C

Darlington, K. (1996). Basic expert systems . Information Technology in Nursing . London .

De Guzman M.P. (2010); Development of a Knowledge-Based Expert System for Intersection Improvement, PhD Dissertation, Institute of Civil Engineering, University of the Philippines, Diliman, Quezon City.

De Guzman M.P. (2011); Application of a Knowledge-Based Expert System to Alleviate Traffic Congestion and Road Crashes at Intersections

D. Keshav, H. Zabeer, and M. A. Hossain, (2005) "Loan Risk Analyzer based on Fuzzy Logic," In Proceedings of the 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service

F.D.R.E Population Census Commission CSA (2011); the 2007 population and housing census Ethiopia Results for Addis Ababa City Administration Statistical report: Addis Ababa 2011 Transport policy of Addis Ababa

Harmon P. & King D. (2005) Expert Systems: Artificial Intelligence in Business**.** Copyright 1985 by John Wiley & Sons, Inc

Hayes-roth F., Waterman D.A., Lenat D.B. (2001): Building expert system: Addison-wesley, Reading, Mass.

Hegyi, B. De Schutter, S. Hoogendoorn, R. Babuska, H. van Zuylen, and H. Schuurman (2001); A fuzzy decision support system for traffic control centers: Intelligent Transportation Systems. Proceedings: 2001 IEEE, pp. 358-363, 2001

Henok. (2011)A case based reasoning knowledge based system for hypertension management. Msc. Thesis, Addis Ababa University, Ethiopia

Hongguo X., Huiyong Z., Fang Z. (2010). Bayesian network-based road traffic accident causality Analysis: In Information Engineering (ICIE), WASE International Conference.

Hutton Pamela, Dougherty M. (1991) Modern artificial intelligence techniques applied to Transport

Ismail N, Ismail A, Atiq R. (2009); an overview of expert systems in pavement management: European Journal of Scientific Research 30(1):99–111.

Kesarwani, P., & Misra, A. (2013). selecting integrated approche for knowledge representation by comparative study of knowledge representation schemes. International Journal of Scientific and Research Publications.

Kingston, John, Nigel Shadbolt, Austin Tate (1996). CommonKADS models for knowledge based planning. In AAAI/IAAI, (1) 477- 482

Krishnaveni S., Hemalatha M (2011). A perspective analysis of traffic accident using data

mining techniques: International Journal of Computer Applications: Published by Foundation of Computer Science, 23, 2011, pp. 40–48

K. Metaxiotis. (2005). Leveraging expert systems technology to improve service industry. European Business Review 17:3, 232-241.

Liebowitz, J. (Series Ed.) (1991), Operational Expert System Applications in the United States, Europe, Pacific Rim, Canada, and Mexico, Pergamon Press, Oxford

Liebowitz, J. (Ed.) (1994a), Worldwide Expert System Activities and Trends, Cognizant Communication Corp., Elmsford, New York, NY.

Liebowitz, J. (Ed.) (1994b), Proceedings of the 2nd World Congress on Expert Systems, cognizant Communication Corp., Elmsford, New York, NY

Liebowitz, J. (1990), the Dynamics of Decision Support and Expert Systems, the Dryden Press, Fort Worth, TX.

Mekasha (2015); road traffic accident controlling mechanisms: Addis Ababa traffic office: Addis Ababa, Addis Ababa University

Morgan A., Mannering F. L.: (1852–1863); The effects of road-surface conditions, age, and gender on driver-injury severities. Accident Analysis & Prevention, 43, 2011, pp

Mossie G. (2009): Applying data mining with decision tree and rule induction techniques to identify. Data mining models for accident severity analysis

Ossenbruggen, P. J., J. Pendharkar, et al. (2001). "Roadway safety in rural and small urbanized areas." Accidents Analysis and Prevention 33(4): 485-498.

Reason, J.; Manstead, A.; Stradling, S.; Baxter, J.; Campbell, K. (**1990**) Errors and violations on the roads: A real distinction?

Regassa Z. (2009) Determining the degree of drivers responsibility for car accident: the case of

Addis Ababa traffic office; Master's thesis, Addis Ababa University: determinant factors of drivers and vehicles in support of reducing and controlling road traffic. Master's thesis, Addis Ababa University

Reynaud, Chantal, et al.(1997). The notion of role in conceptual modeling: Knowledge Acquisition, Modeling and Management: Springer Berlin Heidelberg, (221-236).

Sajja, P. S., & Akerkar, R. (2010). Knowledge-Based Systems for Development. In Advanced Knowledge Based Systems: Model, Applications & Research (pp. 1-11).

Sarma, S. Kr., Singh, Kh. R. & Singh, A. (2009); An Expert System for diagnosis of diseases in Rice Plant: International Journal of Artificial Intelligence, **1**(1):26-31.

Seblewengel, E. (2011). prototype knowledge based system for anxity mental disorders. MSC thesis. Addis Ababa: Addiss Ababa university.

Snie zynski, B.; (2003); Proof searching algorithm for the logic of plausible reasoning: In Klopotek, M., et al. (eds.) Intelligent Information Processing and Web Mining; Advances in Soft Computing

Solomon, G. (2013). a self learning knowledge based system for diagnosis and treatment of diabetes. Addis Ababa: Addis Ababa university.

Speel, P. H., A. Th Schreiber, W. Van Joolingen, G. Van Heijst, and G. J. Beijer.(2001); conceptual modeling for knowledge-based systems. Encyclopedia of Computer Science and Technology

Srinivasan R, Harkey D, Tharpe D (2008) Development of a Web-Based Expert System for setting Speed Limits in Speed Zones.

Stapko, W. (1990), "Knowledge management: a fit with expert tools", Software Management, November

Wentworth J.A. (1990). Developing knowledge-based expert systems, VTT symposium 116, OECD workshop on knowledge-based expert systems in transportation, Vol-1, held in Espoo, Finland, Technical research center of Finland, Espoo.

Wen W. A (2008); dynamic and automatic traffic light control expert system for solving the road congestion problem. Expert Systems with Applications.34 (4):2370–2381.

WHO (2009); the rate of Road traffic accident in Ethiopia: Addis Ababa city

Yeh C, Ritchie S.G. (2003). Schneider J.B. Potential applications of knowledge-based expert systems in transportation planning and engineering.

Appendices Appendix I: Interview questions to Domain Experts

The main objective of this interview questions is to elicit knowledge from road traffic experts that will help for domain know how and the development of a knowledge based expert system for road traffic accident and congestion decision making. The interviewer records the respondents' response using pen, pencil and paper. I thank you in advance for your willingness and valuable time.

1. What are the road traffic accidents?
   Balaan daandii Traafikaa fi Traafiki Nama YKN Konkolaataa maaliidha?
2. What are adapted road traffic problems (accident and congestion) in Jimma city?
   Rakkoon daandii Traafika magaalaa Jimmaa Keessaatti mullatu maaliidha?
3. What are the main risk factors that cause an accident repeatedly?
   Akka balaan uumamuuf wantoonni sababii ta'u danada'an maalfaadha?
4. Are there any predetermined policies (rule and standard) to manage an accident based on their severe?
   Sadarkaa balaa irraati hundaa'uudhaan balaa ta'an to'achuuf heerri YKN qajeelfamni bu'uraaye jiraa?
5. What rule are available to determine the severity of an accident and how to determine it?
   Sadarkaa balaa tilmaamuuf qajeelfamni hordofamu maaliidha?
6. How to determine the probable risk of an accident in future time and what preventive measurements were considered?
   Wantoota balaa qaqqabsiisanii fi tarkaanfiin furmaataa akkamittiin murteeffama?
7. What are the main decisions the traffic officer make to solve road traffic problem?
8. Do you have a standard guideline that you use for the decision making and approval process?
   Murtee kennuu fi mirkaneessuuf ulaagaaleen hordoftan qabduu? Ulaagaa akkamiiti?
9. When, where decision making idea was generated or incase what influential matter?
   Yaadni murtee kennisiisuu yoomii fi eessatti qophaa'a? Maaltus akka murteen kennamuu qabuuf dirqamsiisa?
10. What are the common decision on the accident and congestion?

Murtee dhaabbataan erga balaan ta'ee fi bakka balaan ni ta'a jedhamutti kennaman kan akkamiiti?

11. How you manage the challenges during accident problem and making decision?

    Yeroo rakkoowwan balaa uumamanii fi murtee irratti kennaman wanta isin mudatan akkamiti manage gootu?

Appendix II: Prototype Evaluation form for the Domain Expert

This is an evaluation form to be filled by road traffic experts in order to evaluate the applicability of the KBES for road traffic accident management and control and decision making. I thank you in advance for your willingness and valuable time. Description of the parameter values are as follows.

Performance value	1	2	3	4	5
Description	Poor	Fair	Good	Very good	Excellent

Instruction: An expert will requested to tick on the appropriate value for the corresponding parameter of the KBES for road traffic accident management and control and Decision Making.

Evaluation Parameters	**Performance value**					
	1	**2**	**3**	**4**	**5**	**Average**
Adequacy and clarity of result for decision making						
Relevancy of the retrieved cases in the decision making						
Fitness of the final solution to the problem at hand						
Ease of use of the KBES system						
Relevance of the attributes in representing the road traffic case						
Efficiency of the system in time						
Resource adequacy of the system						
Interactivity of the user interface						
Rate the significance of the system in the domain area						
			Average			

1. How was Adequacy and clarity of result for decision making?

   Bu'aan murtee kennamerraa argamu hagam sirrii fi fudhatamaadha?

   Poor ☐ Fair ☐ Good ☐ V. Good ☐ Excellent ☐

2. How was Relevancy of the retrieved cases in the decision making?

   Murtee kennuurratti ta'iiwwan funaanaman hagam kan systemii wajjin wal fakkaatu?

   Poor ☐ Fair ☐ Good ☐ V. Good ☐ Excellent ☐

3. How exactly Fitness of the final solution to the problem at hand?

   Furmaannii kennamu kan harkaa fi kan systemii hagam wal simata?

   Poor ☐ Fair ☐ Good ☐ V. Good ☐ Excellent ☐

4. How much ease of use of the RTA KBE system?

   Systemii balaa konkolaataa to'achuutiif hojjetame kana gargarmuun hagam salphata?

   Poor ☐ Fair ☐ Good ☐ V. Good ☐ Excellent ☐

5. Is that attributes relevant with Road traffic cases represented?

   Bu'aa ta'iiwwanii kan bakka bu'an hagam bu'aa isaa wajjin wal simatu?

   Poor ☐ Fair ☐ Good ☐ V. Good ☐ Excellent ☐

6. How exactly efficiency of the system in time?

   Yeroo itti gargaaraman hafam salphata?

   Poor ☐ Fair ☐ Good ☐ V. Good ☐ Excellent ☐

7. Is the resource was adequate to the system?

   Systemii kana kan bu'a qabeessa godhan qabeenyi jiruu?

   Poor ☐ Fair ☐ Good ☐ V. Good ☐ Excellent ☐

8. How were interactivity of the user interface?

   Itti gargaraamaan hagam fuula systemii waliin wal bara?

   Poor ☐ Fair ☐ Good ☐ V. Good ☐ Excellent ☐

9. How you rate the significance of the system in the domain area?

   Barbaachisummaa systemii kanaa naannoo keerraatti hagam faayida qaba?

   Poor ☐ Fair ☐ Good ☐ V. Good ☐ Excellent ☐