Predictive Analysis of Pre-Hospital Care Ambulance Services in Western Province - Sri Lanka

P.M.M. Riza

169330 M

Faculty of Information Technology

University of Moratuwa.

July 2019

Predictive Analysis of Pre-Hospital Care Ambulance Services in Western Province - Sri Lanka

By P.M.M. Riza

169330 M

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa, Sri Lanka for the partial fulfillment of the requirement of the Degree of Master of Science in Information Technology

July 2019

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given

Name of Student	Signature of Student
P.M.M. Riza	-----------------------
	Date:
Name of Supervisor	Signature of Supervisor
S.C. Premarathne	-----------------------
	Date:

Dedication

We are dedicating the results of our research and dissertation to decision-makers in the Ministry of Health, Nutrition and Indigenous Medicine who are trying to improve ambulance services for pre-hospital care in Sri Lanka. We also dedicate this system to all those who have generously dedicated their valuable time to advise and carry out this research, especially to my supervisor, Mr. S.C Premarathne. In Sri Lanka, pre-hospital care data analysis has not been carried out effectively with the appropriate techniques. It was with this thought in my mind that we did this research. I hope that the research and results described in this document will provide a useful overview by analyzing pre-hospital data to provide solutions related to pre-hospital issues.

Acknowledgement

It is a great honor for me to take this opportunity to express my gratitude to all those who have helped me throughout this project and guided me to the success of this project.

First of all, I would like to thank my project supervisor, Mr. S.C. Premarathne, who has spent his precious time leading this research to make it a success. Then Dr. M.F.M. Firdhouse, who taught research methodology and project management, was the foundation of this research. Then my next big thank goes to Dr. Rewanika Senawirathne (MO-Planning), a colleague from the headquarters of the Ministry of Health, who guided me in every movement of struggles in terms of health planning and management in this project .

In addition, my thanks go to all the lectures in MSc in Information Technology degree programme of Faculty of Information Technology, who have gone to the trouble of honing our knowledge and ideas throughout these two years, as they were the enlightenment that illuminated our path to success. Finally, I express my deepest gratitude to my wife and daughters who have helped me in all the way by encouraging me to carry out this project successfully.

Abstract

Pre-hospital care is emergency medical care provided to patients before their arrival at the hospital after activation of emergency medical services. This is a crucial element of all emergency care systems, but the needs of emergency patients must also begin early in the field. It traditionally included a range of care from spectator resuscitation to the treatment and transfer of legal emergency medical services. New concepts of care, including Emergency Medical Technicians (EMTs), emergency care practitioners, and physician-provided pre-hospital emergency medicine, are redefining the scope of pre-hospital care. The pre-hospital care data is large and exponentially growing daily in Sri Lanka. There are few researchers who have analyzed pre-hospital data such approaches are not capable of handling big data effectively and not efficient in predicting/describing the issues attached with the pre-hospital data.

Hence the research has been conducted to analyze pre-hospital data efficiently to explore different issues pertaining to the pre-hospital dataset. It is hypothesized that analyzing pre-hospital dataset can be done through data mining according to the output want to achieve through predictive or descriptive techniques. The solution takes the pre-hospital dataset as the input and predicts the factors attached to the particular pre-hospital issues with its associative causes. The overall design of the research consists of two research question, one question used predictive mining based solution and the other one is based on descriptive mining. Prediction in data mining was used to predict the Triage category based on the attribute, the classification and association were used in finding the factors and their relationship that related to the attributes. Finally, the data model analysis using data mining techniques are evaluated for their performance by using accuracy, error rate, and elbow methods, etc. I hope this research approach will be useful for decision-makers and policy-makers in the Ministry of Health, Nutrition & Indigenous Medicine.

Table of Contents

Page

List of Figures

List of Tables

CHAPTER-1

Introduction

1.1 Introduction

Pre-hospital care is an essential component of the trauma system. The Ministry of Health, Nutrition & Indigenous Medicine establishes emergency medical services in Sri Lanka. Intended for the provision of emergency ambulance service, including emergency care and transportation to hospitals. The goal of prehospital care is "During this generation and for future generations, everyone in Sri Lanka will have access to qualified pre-hospital medical personnel, ambulances are available to safely transport the sick and wounded to either Inadequate pre-hospital care is eliminated, so that the medical and nursing staff of the hospitals receives the patients that it is able to professionally treat and rehabilitate the society as contributing citizens.

Several Ambulance services are available at the moment on paid basis especially private hospitals having their Ambulance services in Colombo.

In 2005, with the support of Johantier International from Germany, efforts to establish an EMS system were relaunched. Technical support and four fully-equipped ambulances were provided, with local implementation of the donations by St. John Ambulance Sri Lanka. With this infusion of equipment and training, the first EMS system was launched in August 2005, using the 110 three-digit toll-free telephone number.

The Ministry of Health and Local Government Health Ministries collaborated to implement pre-hospital systems in nine districts outside the capital city of Colombo. Four of the district programs Colombo, Badulla, Galle and, Kandy are fire-based systems. The Anuradhapura and Kurunegala Pre-hospital care systems are based at regional hospitals and the programs in Jaffna and Mannar were implemented using the resources of the regional director of health services. Where the locations were managed by a different set of rules and provided services also in-efficient therefore, the public expected a simply accessible & centralized system with totally free services.

In 2015, India offered a gift to Sri Lanka, as proposed by an emergency medical service, after personal experiences with the failures of the Sri Lankan system as well as urgent requests from doctors. It has also been supported by the Sri Lankan government, allowing the world-renowned GVK Emergency Response Institute to train hundreds of Sri Lankan Emergency Medical Technicians (EMT's) and paramedics, as well as to donate hundreds of Ambulance with the grand of $ 7.5 million to initiate the service in the Western and Southern Provinces of Sri Lanka as a phase-1.

The ambulance service was based on the 108 services of India. Launched in 2016 with 250 highly qualified EMTs, 250 pilots (paramedics) and 50 call center operators. The service derives its name from "1990 SUWASARIYA Ambulance", the toll-free number used to call the service, which is accessible for free from any mobile network, via the command and control center located in Rajagiriya Colombo.

1.2 Current Status

Recent years have seen the modernization and some standardization of practice, an explosion of services able to provide pre-hospital critical care, and the hard-won recognition of the subspecialty of Pre-Hospital Emergency Medicine (PHEM), with dedicated training schemes for PHEM practitioners. The pre-hospital care is a spectrum of a topic which was studied thoroughly and identified the best model with the notification of the Triage category level to implement in Sri Lanka. The Figure-1.1 below describe the Triage Category, Severity Type and the Acuity (Maximum waiting time) of cases. All state sector hospitals are being ready to receive the cases by an Accident & Emergency Unit of the Hospitals running 24 hours based on the triage level [1].

TRIAGE CATEGORIES

TRIAGE CATEGORY	SEVERITY	ACUITY (Maximum waiting time)
Category 1 (Red) Immediate (Resuscitation)	Life threatening	Immediate
Category 2 (orange) Emergency	Imminently Life threatening	10 minutes
Category 3 (Yellow) Urgent	Potentially Life threatening	20 minutes
Category 4 (green) Semi urgent (standard)	Potentially serious	30 minutes

Figure-1.1: The Triage categories & Acuity

The Figure-1.2 below describe the flow of activities after receiving a call and identifying the target location and to dispatch the patient to the nearest state sector hospital after activated the pre-hospital care services.

Required policies have been improved & circulated to all state sector hospitals to receive the patient from the pre-hospital care providers without any issues.

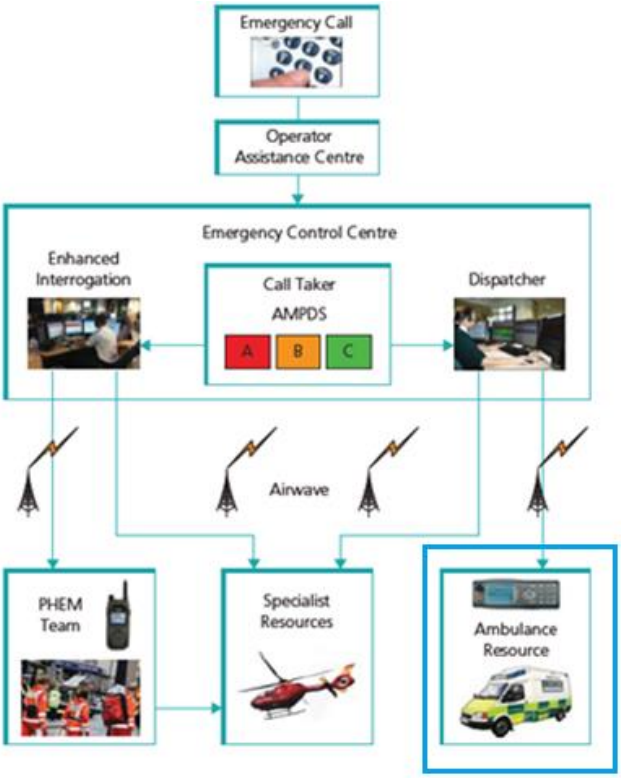


Figure-1.2: The flow of activities after receiving a call

The case of Road Traffic Accident (RTA) handling trauma patient with the required evident base investigation by the Department of Traffic Police and the Insurance providers the activities had been improved significantly in past years.

Emergency location and the ambulance movement are recorded by using the latest GPS technology to enhance the response time with dispatcher time

Monitoring & logging all related activities with high security for the databases have been improved by the IT experts.

1.3 Aim and Objectives

The main objective of this research is to analyze the pre-hospital care ambulance service existing in the Western Province, managed by the Ministry of Health, Nutrition and Indigenous Medicine of Sri Lanka. The following goals and objectives have been identified for this analysis to improve services.

1.3.1 Aim

The goal is to gain in-depth knowledge of data model construction and the application of data extraction techniques to existing pre-hospital care ambulance services in the Western Province of Sri Lanka.

1.3.2 Objectives

- Identify the average number of calls/day, Response time, and Dispatch time.
- Identify the number of Major categories followed by the ambulances.
- Find out the relationships among the attributes for Major categories using the Descriptive analysis, Association and or Clustering.
- Identify the standard of the triage category being used by the ambulance staff and predict the best model to fine-tune the triage level based on the attributes and other facts
- Find out in which District /Area has to be placed more ambulance to handle the requirements of the public.
- How this analysis will help decision-makers to enhance the services and improve the health status of the public.

1.4 Proposed Solution

In this research, to analyze the data, we proposed a method using data mining. Researchers have identified data mining as the best solution for digging and uncovering hidden models in large data repositories. Using software tools support with its ability to simultaneously handle a large number of dynamic variables.

Initially, the prehospital care dataset was pre-processed and prepared for data analysis using Rapid Miner. Data mining is about discovering correlations, patterns, trends, or relationships by looking for a large number of datasets. It is selected as the basic methodology for this research. Using clustering and association techniques, the evidence of the relationship between the attributes was identified (Critical Condition = yes). We then analyzed a descriptive analysis based on the relationship between the attributes needed to solve the problem of under-research in the first step of the data mining process.

As a solution to the sub-search question, the classification method used to classify the dataset to predict triage categories using the Rapid Miner. Where the triage category depends on the severity and acuity time. Instead of assigning the manual triage category, the triage category predicted above will improve the patient's delivery to the hospital by the ambulance staff as required to protect the patient as well as service delivery under the service level agreement (SLA).

1.5 Structure of the Thesis

The thesis is organized in the following order. Chapter 2 presents similar work done by others while Chapter 3 explains how the technology adopted in this domain. Chapter 4 describes the approach of the research and Chapter 5 illustrates the analysis and design. Chapter 6 concerns the implementation of the analysis. Chapter 7 is an evaluation. Finally, Chapter 8 presents the conclusion and list of references at the end of the section.

CHAPTER-2

Review of others' work:

There are many types of research available in pre-hospital care, when we check worldwide there are many research documents describe the activities are as follows.

2.1 Emergency Care in International

When we consider a developed country like Europe, according to R. Fairhurst's, the provision of pre-hospital care varies between the 25 European countries, where the provision of care is not uniform in the same way as the service provider, the system, the number of the hot-line and quality of services. In most countries, pre-hospital care is provided by physicians, often as part of training in another specialty. In England, there are 31 ambulance services with different procedures and equipment. Who make confusion and difficulties for the public to contact in case of an emergency those who move from time to time within the region [2].

France is the only country where prehospital care is recognized as a specialty, where they are included in the recently recognized Emergency Medicine. Indeed, only the United Kingdom, Ireland and France recognize emergency medicine at the hospital. The disadvantage is that in England, 31 ambulance service providers provide pre-hospital services with different standard operating procedures and equipment. There are three ambulance service systems and three telephone consulting software systems in practice. To avoid such problems, the Sri Lankan system has put in place a unique set of rules and procedures, with the unique tall-free number 1990 from any mobile network available in islandwide. The implementation of joint activities and the exchange of information between national and regional organizations on the pre-hospital management of medical emergencies throughout the European Union is the advantage of these systems.

Indonesia is a developing country that until recently did not have a public ambulance service with a prehospital care system. E Pitt, and Pusponegoro, states that Emergency Hotline 118 is the only public ambulance service in Indonesia. It's not funded by the government, but a fee for those who can afford to pay. To access it, dial the toll-free number 118. The Jakarta call center receives around 50 to 75 calls a day.

In a city of 10 to 12 million people, it shows how much service has yet to go to be used universally[3]. The service has been upgraded recently with 119 hotlines with the facility upgraded, but the majority of people use private vehicles to transport patients to the hospital due to issues related to these services [4].

The disadvantage of the Indonesian system is that this service is not free anymore and the public can hire vehicles to the hospital as soon as possible with less amount compare with the price of the ambulance service providers. The service covers the whole of Jakarta (Area of 661 square km with a population of 10 to 12 million), with only 26 ambulances and 12 motorcycles. They cannot cover such a large crowded area with few ambulances to cope with the urgent needs of the public. Ambulances are located in 10 strategic locations around the north and center of the city, which is much surprising because of response times are very poor due to lack of vehicles and congestions on the road. The advantage in Sri Lankan system is that both the entire pre-hospital service and indoor-hospital services are free for all the public so anyone can get the service free of charge at any time which will enhance the healthy life of the public.

Another researcher in Malaysia by Aliza Sarlan1; et al [5] states that there should be an effective pre-hospital emergency notification system to allow for communication between the emergency medical team and the hospital: there is no proper channel of communication between these two parties regarding the care of the patient. Therefore, the author focuses on the pre-hospital emergency notification system, which consists of an Android mobile application that sends the medical severity of road accident victims to the hospital through a web-based system. The end-user of the hospital received the post notifications and information about the victims in time for pre-hospital arrangements to receive the patient for further care.

The author is focusing very efficient system in Malaysia which is very advanced in the pre-hospital setup that interacts both the parties of the emergency medical team and hospital staff to work together for further arrangements and care of the arriving emergency patient to the hospital. Where both parties should work in one common web-based Hospital Management System / Patient Management System which provides the platform to enhance the services. The Sri Lankan system also has the disadvantage of the above setup where the ambulance staff utilizing very advanced web-based Patient Management System but on the other hand, the Government hospitals are still running with the paper-based system in islandwide. Also, the author proposes the improvement of systems so that the medical records of each patient are highly private. The future development of the system should focus on security. All data transfers must be encrypted with strong cryptography to prevent any hacking activity. The electronic health record database should only allow authorized users to access it.

Prevention is a fundamental value of any health system. Nevertheless, many health problems will continue to occur despite prevention services. J.A. Razzak and A.L Kellermann report that in developing countries, diseases and time-related trauma, such as serious infections, hypoxia caused by respiratory infections, dehydration caused by diarrhea, intentional and unintentional injuries, postpartum bleeding, and acute myocardial infarction. The provision of timely treatment in life-threatening emergencies is not a priority for many health systems in developing countries. This article examines evidence indicating the need to develop and/or strengthen emergency medical care systems in these countries [6].

The author focuses on developing countries, health care generally does not focus on emergency medical care. Although health promotion and the prevention of illness and injury are core values of any health system, many acute health issues will continue to arise. Integrating a basic level of emergency medical care into health care systems could have a significant impact on the well-being of populations. This would respond to the self-perceived needs of the population and reduce the long-term human and economic costs of illness and injury.

Priority should be given to developing minimum guidelines for emergency medical care in low-income countries. The effectiveness of this care could be evaluated by implementing pilot programs in several low- and middle-income countries. This would help determine the extent to which emergency medical care systems save lives and at what cost. The Ministry of Health care Nutrition and Indigenous Medicine in Sri Lanka has already implemented the policy-based priority and Accidental & Emergency care Policies in practice for all Government and Private health care institutions in islandwide.

2.2 Emergency Medical Services Policies & The status in Sri Lanka

Accident and Emergency (A&E) care is a challenging and complex area of practice in Sri Lanka, presenting a variety of challenges for patient care. In recent years, the incidence of accidents and emergencies has increased. The Ministry of Health, Nutrition and Indigenous Medicine has identified the need for a strong accident and emergency care policy to develop comprehensive A&E health care services as a priority basis in public hospitals of Sri Lanka. The guidelines described in this document include A&E's operational structure and care model, triage system, infrastructure development, quality improvement, the standard for human resource requirements, and A&E information system [7].

Sri Lanka provides free health care to its people and adheres to the principles of social justice, equity and respect for human rights. The guiding principles of the national accident and emergency policy are:

a. Protection of the right to health and value for life
b. Equity, social justice, and cultural adequacy
c. Patient-centered care
d. Multidisciplinary Approaches to Comprehensive Care
e. Efficiency and effectiveness
f. Technical quality and service
g. Affordability and sustainability
h. Continuity of care
i. Respond to emerging health needs through evidence-based approaches

The implementation of the policy will be based on the national strategic framework for accidents and emergencies and on the implementation guidelines. The strategic framework describes the proposed activities for each of the strategic objectives, and the implementation guidelines refer to the following areas:

a. Operational Structure and A & E Care Model
b. Sorting system for A & E units
c. Infrastructure Development Directive
d. Standard human resource requirements for engineering and evaluation units

e. Standard Medical Equipment Requirements for A & E Units
f. Standard equipment, facilities and capacity building required for ambulances
g. List of standard drugs for a unit of medicine and assessment
h. Standard resuscitation cart for A & E units
i. Information system for A & E units
j. Capacity building of human resources within A & E units
k. Quality improvement in A & E units

Annual operational plans will be developed for each of the above strategic areas based on implementation guidelines for each level of A & E care.

It is also recognized that a significant portion of the burden of injury can be avoided if evidence-based policies are in place and the relevant programs are implemented. A national policy is essential to give appropriate priority to injury prevention and management, to efficiently organize resources and to coordinate & guide the different sectors involved in injury prevention. Because of these facts, the Ministry of Health has developed the National Injury Prevention and Management Policy. [8]. This strategic framework will address the following areas:

1. Primary Prevention - Reducing Risk Exposure and Injury Prevention by Adopting Safer Behaviors and Environments
2. Secondary prevention - In case of injury, reduction of the severity of the injury and its impact, for example. Early diagnosis and appropriate management of an injury by applying basic first aid at the scene of an incident and ensuring prompt transportation to the hospital in an ambulance with facilities to prevent injury has more serious consequences.
3. Tertiary prevention - Reduction of the consequences of injuries through post-event care (eg. emergency care, essential trauma care, physical and psychological rehabilitation) - involves the prevention of new complications in the form of more serious injuries, disability or death.

National policy will cover unintentional injuries: traffic accidents, drowning, burns/fires, falls and poisonings. Intentional injuries are dealt with in the mental health policy. Secondary prevention and some aspects of tertiary prevention are considered common to all types of injuries and this policy concerns the required service structures. Disaster preparedness policy is also consistent with this policy of service delivery.

2.3 Emergency Care with Information Technology

Information technology is involved in almost all areas of Emergency Medicine to simplify the process and the task. The author, K. Wimalaratne, IL Lee, Lee KH, HY Lee, Lee JH and IH Kang, describes the emergency medical service systems (EMS) in Sri Lanka: past problems, challenges of the future, in terms of development, delivery and future ideas for the EMS in Sri Lanka [9].

The author published these documents in early 2012 at that time there is no qualified emergency Physician in Sri Lanka. However, a specialist training program for emergency physicians was launched. Nowadays, the Physicians are in place in most of the hospitals in Sri Lanka, there is no formal system for training emergency medical technicians (EMTs). Therefore, The Ministry of Health Nutrition and Indigenous Medicine has expected the right source to provide the training program for Sri Lankans in-order to establish the Emergency Medical Services. By the help of Indian Government, the Ministry has initiated the training program for thousands of Sri Lankans and launched the services in 2016 as “1990 Pre-hospital care Ambulance Services”. Sri Lankans generally use taxis or their private vehicles to get to the hospital in an emergency situation before starting these services. Also, the author pointed out that Sri Lanka has a long way to go to provide integrated, effective and functional prehospital care. First, the country needs a centralized communication center with properly trained regulators. Such a resource would communicate with various emergency services, including fire brigades and the police, to provide support in defining civilian relief needs, massive casualties, and response activities in the event of an emergency.

Emergency medical services cover a wide area to provide the facility for those in need of medical treatment or to bring patients to final care. In overcrowded cities, rapid access to emergency care becomes difficult as the number of people increases. The author, A. Akram, followed by at el, focused on an end-to-end communication solution in the form of an integrated platform with applications that can help patients and rescue service providers in critical situations, effectively and efficiently [10]. The advantage of this system is an end-to-end solution is proposed to help patients; emergency service providers and physicians communicate effectively in emergency situations. The proposed platform used current mobile and web technologies to automate the existing emergency system with a new approach. The goal of the Savior of Life was to minimize response time and provide reliable information that can help people in emergency situations. With this system, ambulances can be routed to emergency points efficiently and continuously to the hospital.

The technological improvements have played an influential role in improving the quality of health services and have introduced significant efficiency in EMS and a reduction in the mortality rate [11]. This article is expressing the effectiveness of emergency care has long been expected, and it is becoming increasingly distinctive with increasing population, poverty, crime rate, and threats such as terrorism. Pre-hospital and emergency care in Pakistan is far from meeting the standards of emergency care around the world. Basic infrastructure and human resources are available; however, technology initiatives are needed to equip ambulance service providers and hospitals with innovative ICT solutions to improve the efficiency of emergency care services. The point "ICT solutions to improve the efficiency of emergency care services" is the critical point of this article which has to be considered about the Hospital Management System / Patient Management System in a centralized manner in Sri Lanka. At present, the Government Hospitals are running without an efficient Hospital Management System or Patient Management System which leads to the communication barriers of the pre-hospital care management system to update the patient status of the admitted patients in the hospitals by the ambulances.

EMS and pre-hospital care are difficult to assess. Therefore, the real efficiency and value of such systems are difficult to define. Global indicators are very difficult to develop because of the multitude of variations and the merging of factors involved in this area [12]. When considering "pre-hospital care", particularly from the point of view of medical services, it may be useful to consider services as belonging to one of the following categories;

a. Services in which the pre-hospital period is relatively short and the host institution is a sophisticated trauma and emergency center. This is the situation in many western urban centers.

b. Services in which the length of the pre-hospital period is relatively long in a sophisticated center. In this situation, rural services may have to travel long distances.
c. Services in which the pre-hospital period is relatively short, but in a small, unsophisticated trauma and emergency center. This can happen in rural areas.
d. Services in which the pre-hospital period is relatively long in a small, unsophisticated reception center. Pre-hospital ALS personnel can play a role in strengthening emergency department resuscitation teams in these circumstances.

These four scenarios have very different considerations and implications for the activity, practice, and supervision of the EMS. They could form the basis of an appropriate classification for a more accurate and realistic assessment of EMS systems. This would help to develop more appropriate indicators adapted to different types of services. The author is indicating highly valuable points in relation with the analysis of the EMS system wherein Sri Lankan system falls under the category of ‘c’ that the pre-hospital period is relatively short but in a small, unsophisticated trauma and emergency center. Due to inefficient EMS system running in Sri Lankan Government hospitals therefore, the individual attention needed in this area to uplift the sophisticated systems in A&E units of the hospitals.

2.4 Pre-Hospital Trauma Care

The author S.D. Dharmarathne, A.U. Jayatilleke, and A.C. Jayatilleke state that road crashes, injuries, and deaths in Sri Lanka all increased between 1938 and 2013. However, there were fluctuations during this period; and several factors are associated with these changes. Fluctuations over time in road traffic accidents and crashes in Sri Lanka are associated with changes in Politics, Economic status, and Road traffic circulation [13]. As in Sri Lanka, these increases have been associated with an increase in the number of vehicles, week enforcement of the traffic law and underdeveloped road infrastructure. The burden of road accidents in Sri Lanka could be reduced through better enforcement of the Highway Code, restrictions on the importation of two or three-wheel motor vehicles and the introduction of new policies to improve road safety.

In early days the road traffic accident casualties were carried out by the public and the casualties were transferred to the hospitals through personnel vehicles without implement any type of pre-hospital care. After implementing this “1990 services” the cases were handled in an advanced manner in order to care the casualties on scan and to the transportation to the hospitals for further care.

The key elements of successful pre-hospital trauma care are the well-known ABCs of trauma care: Airways, Breathing, and Circulation. The establishment and safety of the airways, ventilation, fluid resuscitation, as well as rapid transportation to the most appropriate trauma center, represent the pillars of trauma treatment in the field. Although the ABC in Trauma Care has not been questioned or altered, new technical tools and procedures have been developed to help the pre-hospital care provider achieve these goals in the pre-hospital setting and thereby improve the outcomes of the patients [14]. In order to make sure the Airways, Breathing, and Circulation of the emergency patient the “1990 ambulance” have already built with the necessary equipment with the trained Emergency Medical Technicians are in place to activate the pre-hospital services and to transport to the hospital which is an advantage in existing pre-hospital care ambulance service in Sri Lanka.

2.5 Pre-Hospital care Data Analysis

In this domain, few of pre-hospital care documents were found online in different categories. The author Davis, James S at el also reviewed different types of data analysis, including all the trauma deaths in 2011 that were not carried to a hospital (i.e., died on the scene) or dead on their arrival. Age, Sex, Date of death, Mechanism and list of injuries were recorded. An expert panel reviewed each case to determine the primary cause of death and whether the patient's death was due to injuries that could lead to survival or non-recoverable injuries [15]. The author is used expert penal review method as a human this can be right or wrong which can be arguable for some instances. In my research, I have chosen the best data mining methods and techniques to analyze the pre-hospital data which can produce the solid output for decision-makers and there are no possibilities for conflicting any outputs or decisions.

The author M. Bigdeli, D. Khorasani-Zavareh, and R. Mohammadi states that a cross-sectional retrospective study was designed at different intervals for the pre-hospital care of the RTI's identified at the Ambulance dispatch center in Urmia, Iran, from March 20, 2005, to March 20, 2007. All cases resulted in Ambulance shipments were reviewed and analyzed [16].

Here the author has used Cross-sectional studies which involve in data collected at a defined time. They are often used to assess the prevalence of acute or chronic conditions, but cannot be used to answer questions about the causes of the disease or the results of an intervention. The author has selected an in-efficient method for analyzing the huge amount of data which the methods will not provide the answers for questions about the causes of the disease or the results of an intervention.

Author Seymour, Christopher at el, analyzed the incidence of severe sepsis in pre-hospital emergency care, where standard statistical methods were used to analyze the data to complete the study [17]. In a large community-based cohort of 10 years, the author demonstrates that emergency medical services staff frequently meet patients hospitalized for severe sepsis. The EMS system, as a whole, carries up to 40% of all severe sepsis-related hospitalizations in the emergency department and delivers prehospital care for nearly one hour in the sickest patients. Although most cases of severe sepsis were diagnosed at hospitalization, pre-hospital interventions, including intravenous access, were rare. These results, coupled with the considerable time allotted to prehospital care, point to a significant and growing opportunity to recognize and potentially treat severe sepsis prior to hospital arrival.

Author S. Trauma explained the hospital and emergency health interventions provide rich sources of high fidelity data. Storage, management, and analysis of this data go beyond traditional means and calls for "Big Data" approaches. With non-invasive ambient data sensors and reliable data collection techniques, fractional information from heterogeneous data sources can be assembled in real-time and applied to specific studies. They can also integrate in-depth customer knowledge into the automated learning process. The analysis of characteristics derived from the 'Photopletysmogram' waveform illustrates the benefits of using large data for early prediction of trauma outcome and autonomic resuscitation. [18]. Where the researcher collected massive clinical digital data regularly by high-throughput biomedical devices provides opportunities and challenges for optimal use. This article explains how such data is used in learning prediction models in level 1 trauma centers to aid decision making in traumatized patients.

2.6 Summary of Literature

The prehospital care data is big data, huge in volume updating on every minute available online. For the big data analysis, there is no proper analyzed method used in the literature for online analytical methods. The following chapters describe the appropriate methods for data analysis in this study.

CHAPTER-3

Technology adapted

3.1 Introduction

Chapter 2 examined existing methods and technics in pre-hospital care services to identify issues related to pre-hospital care. This chapter introduces the data mining technology that was selected to analyze data from the Ministry of Health, Nutrition and Indigenous Medicine in Sri Lanka. This chapter highlights the effectiveness of the selected technologies that distinguish it from the technologies applied in the existing literature.

.

3.2 What is Data Mining?

Data mining is an automated process of sorting huge data sets to identify trends and patterns and build relationships, solve business problems, or generate new opportunities through analysis of the Data.

The term "data mining" is used quite widely in the computer industry. It often applied to a variety of large-scale data processing activities, such as data collection, extraction, storage, and analysis. It is not enough to look at the data to see what happened in the past and be able to act intelligently in the present. Data mining tools and techniques allow you to predict what will happen in the future and act accordingly to take advantage of future trends.

Data mining is an essential part of the knowledge discovery process. In this process, the process can include the following steps: data selection, data cleanup, data transformation, model search (data mining), presentation search, interpretation search, and evaluation. Data mining and KDD are often used interchangeably, as data mining is the key element of the KDD process. The term Knowledge Discovery in Databases or KDD, in short, refers to the general process of searching for knowledge in data and emphasizes the "high-level" application of particular data mining methods. It interests researchers in machine learning, pattern recognition, databases, statistics, artificial intelligence, knowledge acquisition for expert systems and data visualization. The unifying goal of the KDD process is to extract knowledge from data in the context of large databases. To do this, it uses data mining methods (algorithms) to extract (identity) what is considered knowledge, in accordance with measurement and threshold specifications, using a database as well as all processing, pre-sampling, and transformation required from this database. This extraction process is performed in Multiple steps as shown in Figure 3.2.1 below.

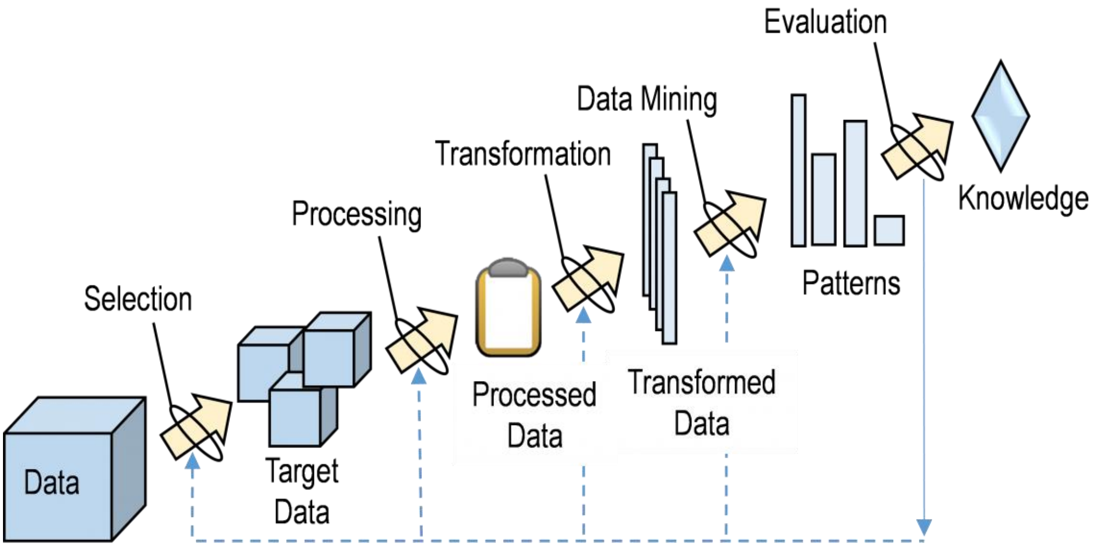


Figure-3.2.1: Discovering knowledge in databases

Here is the list of steps in the knowledge discovery process -

- **Data cleaning** - In this step, noise and inconsistent data are deleted.
- **Data Integration** - In this step, multiple data sources are combined.
- **Data Selection** - In this step, the relevant data for the analysis task is extracted from the database.
- **Data Transformation** - In this step, data is transformed or consolidated into appropriate forms for retrieval by performing summarization or aggregation operations.
- **Data mining** - In this step, intelligent methods are applied to retrieve data models.

➢ **Model Evaluation** - In this step, the data models are evaluated.

➢ **Knowledge Presentation** - In this step, knowledge is represented.

3.3. Data Mining Tasks

There are a number of data extraction tasks such as classification, forecasting, time series analysis, association, clustering, synthesis, and so on. All of these tasks are predictive data mining tasks or descriptive data mining tasks. A data mining system can perform one or more of the above tasks as part of data mining. Data mining tasks can typically be categorized into two types based on the objectives sought by a specific task. These two categories are descriptive tasks and predictive tasks as described below in Figure 3.3.1. Whereas the descriptive data mining tasks illustrate the general properties of the data, while predictive data mining tasks infer the available data set to predict the behavior of a new data set.

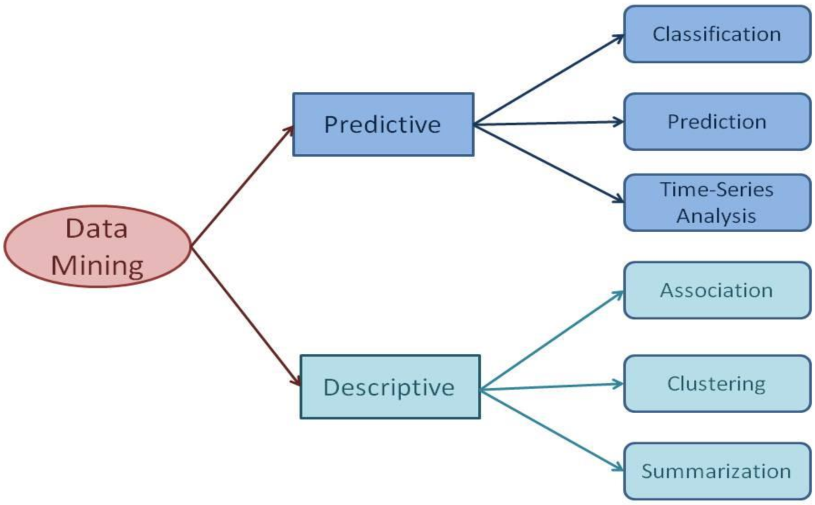


Figure-3.3.1: Datamining Task

Predictive data mining tasks develop a model from the available dataset that is useful for predicting unknown or future values of another dataset of interest. For example, A

Medical Consultant trying to diagnose a disease based on the results of a patient's medical records can be considered a predictive data mining task.

Descriptive data mining tasks typically find data describing models and provide important new information from the available dataset. A retailer trying to identify products purchased together can be considered a descriptive data mining task.

3.3.1. Association

The association discovers the association or connection between a set of elements. The association identifies the relationships between the objects. Association Analysis is used for health care management, product management, advertising, catalog design, direct marketing, and more.

To find out the relationship among the attributes of Incident type, Triage category, District, Sex and Critical requirements were divided into multiple attributes as described in Figure: 3.3.1.1 below.

Incident_type	TRIAGE_CATEGORY	District	Sex	Critical_requirments
Inci_Accident	Tri_Red	DT_Colombo	S_Male	CR_Yes
Inci_Alzheimers_disease	Tri_Orange	DT_Kalutara	S_Female	CR_No
Inci_Baby delivery	Tri_Yellow	DT_Gampaha		
Inci_Burn Injuries	Tri_Green			
Inci_Cancer				
Inci_Chronic kidney disea				
Inci_Chronic respiratory				
Inci_Diabetes				
Inci_Heart disease				
Inci_Osteoporosis				
Inci_Stroke				

Figure-3.3.1.1: Description of categories for Association process

If we can identify the critical requirements associated with which color of the triage category then we can go for further analysis in order to achieve the objectives of this research.

3.3.2. Clustering

Clustering is a data mining technique that creates a useful or meaningful cluster of objects with similar characteristics using the automatic technique. The grouping technique defines classes and places objects in each class, while in classification techniques, objects are assigned to predefined classes. For this pre-hospital care analysis, the clustering techniques were applied to find out the group of attributes which are highly related to the "Critical condition =Yes". According to the result and conclusions then conducted further analysis for the decision making the process of this analysis.

3.3.3. Classification

The classification derives a model to determine the class of an object according to its attributes. The classification consists of assigning cases to categories based on a predictable attribute. Each case contains a set of attributes, one of which is the class attribute (predictable attribute). The task requires finding a model describing the class attribute based on the input attributes. To form a classification model, for the prehospital care dataset, we need to know the class value of the incoming cases in the learning dataset, which are usually historical data, where the class attribute is the triage category. Data mining algorithms that require learning of a target are considered supervised algorithms. Typical classification algorithms include decision trees, KNN, and Naïve Bayes.

Hence, we have conducted a classification process for the data set and predicted the actual triage category based on the other factors available in the dataset which can be utilized for the enhancement of the pre-hospital care ambulance services.

3.4. Reasons for using data mining for pre-hospital care analysis

The health background is generally perceived as "information-rich" and "poor in knowledge". There is a wealth of data available in health systems. However, effective analytical tools to uncover hidden relationships and trends in the data are lacking. Knowledge discovery and data mining have found many applications in the commercial and scientific fields. Valuable knowledge can be discovered through the application of data mining techniques in the health system.

Providing quality services at affordable costs is a major challenge for healthcare organizations. Quality service involves properly diagnosing patients' problems and administering effective treatments. Misdiagnosis / poor clinical decisions can have disastrous and therefore unacceptable consequences.

Data mining technology offers a user-oriented approach to new and hidden data models. Data associated with pre-hospital care is highly confidential due to patient data. The patterns are complex and large. Health administrators can use the knowledge discovered to improve the quality of service. The physician may also use the knowledge discovered to reduce the number of adverse drug reactions and suggest less expensive therapeutically equivalent alternatives. Anticipating future patient behavior over a given history is one of the important applications of data mining techniques that can be used in health care management.

3.5. Summary

Different data mining tasks are at the heart of the data mining process. The Prediction, Classification, and Association data mining tasks are actually retrieved from the required information from the datasets to make the major decisions based on the analysis. The following chapters will follow the approaches for the steps of the decisions.

CHAPTER-4
Approach for Pre-hospital care Ambulance Services

4.1. Introduction

The chapter-3 discussed the technology for analyzing pre-hospital care ambulance services. This chapter presents our approach to analyze pre-hospital care services in detail using data mining under several headings namely hypothesis, input, output, process, users, and futures. This chapter highlights the key futures that distinguish our approach from the existing approaches for pre-hospital care analysis in Sri Lanka.

4.2. Hypothesis

The use of data mining techniques can help predict or explore problems related to trauma cases. Predictive data mining can be used to predict factors affecting different trauma cases. Descriptive data mining can be used to explore the current state of affairs demonstrated in Non-communicable diseases (NCDs).

4.3 Input

Data obtained from pre-hospital care services managed by the Ministry of Health, Nutrition and Indigenous Medicine is used as the initial input for this process.

4.4 Output

As the output of this process, different data patterns related to pre-hospital care can be revealed according to the research question identified. The prediction will be given as output for the research question attached to the predictive task. Summarization will be given for the research question attached to the descriptive tasks.

4.5 Process

In this process of pre-hospital ambulance service analysis with data mining, the standard steps of the discovering knowledge in databases will be established. Throughout the Knowledge Discovery Database Process (KDD Process), the dataset is cleaned, formatted and prepared for exploration and interpretation.

4.5.1 Data Selection

Injuries and other chronic diseases such as cardiac arrest, stroke, sepsis, and obstetric emergencies are important factors in premature mortality and disability in Sri Lanka. The majority of premature deaths due to such urgent conditions are the result of inadequate prehospital care, lack of transportation, or both. Patients may need to be transported as soon as possible to reach a hospital for better care.

Emergency Medical Services (EMS), which may include local, regional or island-wide pre-hospital care systems, play a key role in improving the outcomes of acute illnesses and acute exacerbations chronic diseases. Evidence shows that lack of pre-hospital care negatively affects outcomes of medical, obstetric and pediatric emergencies; the availability of pre-hospital care results in a significant reduction in injury-only mortality, with a greater cumulative effect when safe transportation is associated with rapid emergency care delivered in facilities. As a result, we used the 1990 pre-hospital ambulance service data set for this analysis.

4.5.2 Data Preprocessing

For this research, the 1990 ambulance service data is used as the major data source which contains the Districts of Colombo, Gampaha, and Kalutara for the patient data with the Date, Time, Incident-locations, type of Incidents, etc. The accuracy, completeness, consistency, timeliness, credibility, value-added, interpretability and accessibility are some of the characteristics that research data should have in order to draw well-accepted conclusions.

Incomplete data may come from a data value not applicable at collection by the human errors, Hardware or Software issues, and different considerations between when the data was collected and when it was analyzed. Noise in the data is another problem that reduces the quality of the data. Noisy data can come from human or computer error when entering and transmitting data. Another problem is that of inconsistent data that may come from a functional dependency violation in the linked data. Duplicate records also require data cleanup, resulting in poor data quality. Therefore, the ambulance service data set for pre-hospital care is pre-processed before further analysis.

4.5.3 Data Transformation

This is the stage where data is transformed or consolidated into forms appropriate for extraction by performing operations such as summarization and aggregation. Due to the availability of a huge amount of data and the immense need to adjust them to useful information and knowledge to support government decisions. Smoothing, aggregation, generalization, and normalization are strategies used in the data transformation process.

Smoothing is used to suppress the noise of the data. Aggregation is used for the synthesis and construction of data cubes. Generalization is used to conceptualize hierarchy ascension and normalization to scale to a specified small range. According to the data selected for the secondary research question, appropriate smoothing and aggregations are performed.

4.5.4 Data Mining

Data mining is an essential part of this research that used intelligent methods to extract data models and discovers interesting knowledge. The association is described as a relationship between a particular element of a data transaction and other elements of the same transaction that can predict patterns. The classification also describes the methods used to learn different functions that map each element of the selected data into a predefined set of classes.

4.5.5. Evaluation / Interpretation

The interpretation of the data is the most essential aspect of the knowledge of the extracted data. Two issues are critical: Business Value Recognition from the knowledge models discovered at the data mining stage. The other problem is what techniques or visualization tools should be used to display the results of data mining. Determining the business value from discovered knowledge models is similar to the puzzle solution because different techniques can be used for the same data set with different algorithms. Therefore, the evaluation of mental patterns must be done according to purpose or goals in order to maximize effectiveness. In order to correctly interpret knowledge patterns, it is important to choose the appropriate visualization tool. Many packages and visualization tools are available nowadays, including pie charts, histograms, charts, trees, and distribution networks.

4.5.6. Features

The solution proposed by this research can be used to analyze a considerable volume of data sets on pre-hospital care which come in different forms in a consistent way. With the predictive exploration task, this solution predicts the future instance based on the other attributes available in the dataset. In addition, descriptive tasks describe the model described by the dataset. This solution provides the result by dynamically extracting previously unknown models.

4.6 Summary

This chapter presented our approach to analyzing the pre-hospital care ambulance services data to enhance the existing services in the right direction by using the data mining task. The next chapter shows the design of the approach presented here.

CHAPTER-5

Analysis and Design

5.1 Introduction

Chapter 4 presents the pre-hospital care data analysis approach to identify issues related to pre-hospital care services. This chapter develops the approach and focuses on high-level design and sub-modules in design.

The patient data is highly confidential and it has to be protected in all the way, therefore, In order to avoid the ethical clearance for the patient data, the similar dataset has been generated as per the requirements to carry out this research.

5.2 Research Design

The fundamental purpose of science is to explain a natural phenomenon called theory. In the scientific approach, instead of trying to explain each distinct behavior, the researcher seeks a general explanation that encompasses and connects several behaviors. Therefore, scientific research is an experimental, systematic, controlled, public, and critical investigation of natural phenomena. It is guided by theory and assumptions about the alleged relationships between such phenomena.

In this research of analyzing pre-hospital care data using data mining, we wish to discover how can predict the best triage category for patients at the time of emergency occurrence. Hence, we have to carry out a quantitative data analysis using Experimental Research Design: in analyzing pre-hospital care data we need systematic analysis and investigation of patient status through a controlled, observation and scientific experimentation to draw the conclusions.

In addition, we would like to discover solutions to problems related to the improvement of pre-hospital care services in Western Province. Therefore, we need to perform **quantitative** data analysis using a **descriptive research design.**

5.3 Top Level of Research Design

In this research of analyzing pre-hospital care ambulance service, the data is analyzed by using the Rapid Miner. We would like to discover how can predict the best triage category for patients at the time of emergency occurrence in particular locations and the trends of the major category for the dataset.

Figure -5.2.1 below describe the flow of activities for the analysis. After preprocessing the dataset the target role will be identified as triage category and then followed the steps of Experimental Design to determine the relationship between cause and effect of a situation for the patient triage category of this dataset.

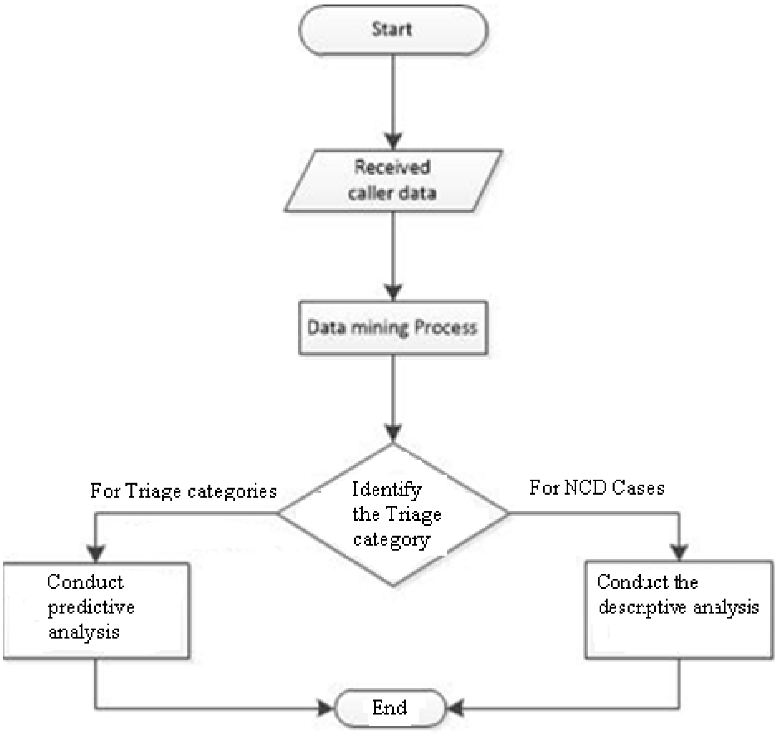


Figure-5.2.1: Flow of activities for the research process.

This is a causal research model where the effect of the independent variable (date, time, Location, Major category, subcategory, etc.) on the dependent variable (Triage category) is observed. This is a highly practical search design method as it helps to solve a problem. The independent variables are manipulated to monitor the changes they have on the dependent variable. The predictive analysis used to predict the best cases for predicted sort category values based on attributes available in this dataset.

On the other hand, a descriptive analysis will be carried out for the best cases in order to find the solutions pertaining to the problems related to the improvement of pre-hospital care services in the Western Province by describing the situation or cases in pre-hospital care data set.

5.4 Detailed Design of the Research

Analyzing pre-hospital care data to find out the issues related to the ambulance staff is identified as the primary research question in this research. According to the dataset, the below Table: 1 describes the logical part of the categories which needed to the descriptive analysis for the sub research questions below.

Incident type	**Triage Category**	**District**	**Time Interval**	**Age Group**
Stroke	Red	Colombo	06:00-12:00 Noon	Up to 16
Accident	Orange	Gampaha	12:00-03:00 PM	Above 16
Baby delivery	Yellow	Kalutara	03:00-09:00 PM	
Burn Injuries	Green			
Cancer				
Alzheimer's disease				
Chronic respiratory				
Diabetes				
Heart disease				
Osteoporosis				
Chronic kidney disease				

Table 1: Logical Description of the Categories

We obtained a sample of “1990 pre-hospital care ambulance service date set” for prediction analysis which was Generated (5000 data). For each data containing detail information on Incident Type, Triage category, District, Age Group and etc. The dataset having multiple attributes among them the above variables logically generate the descriptive information with the other factors for decision-makers the details are as below.

Incident type: The dataset having 11 categories of incident types wherein data set namely: Cancer (487 patient), Diabetes (479), Accident (472), Baby delivery (468), Burn Injuries (468), Alzheimer's disease (462), Heart disease (455), Stroke (449), Chronic kidney diseases (438), Osteoporosis (413). These data can be mapped with the other factors and checked the status.

Triage category: The hospital A&E Unit with the pre-hospital systems are working together for the triage category of Red (103), Orange(1135), Yellow (3143) and Green (619) in this dataset based on the A&E policy instructed by the Ministry of Health, Nutrition and Indigenous Medicine. These data can be mapped with the other factors and checked the status.

District: The Western Province having 3 District Namely: Colombo-The capital city (1289 Cases), Kalutara (1426 cases), and Gampaha (2285 cases) were saved. These data can be mapped with the other factors and checked the status.

Time interval: The service is available 24x7 around the clock on each day. In order to analyze the peak time trends of the Western Province, the timeslots were divided into 3 groups namely: [06:00 AM -12:00 Noon], [12:00 Noon -03:00 PM], [03:00 PM-09:00 PM]. These data can be mapped with the other factors and checked the status / Trends of the cases.

Age group: The hospital admissions depend on the age groups of the patients for better care of services. Where basically age divided into 2 groups which are Pediatrics (Up to age 16) and above 16 in age. The data set is filtered for each group separately and mapped with the other factors and checked the status / Trends of the cases.

5.4.1 Primary Research Question

There is no research carried out in Sri Lanka for the pre-hospital care data to analyze the issues pertaining to the ambulance staff. Hence predominantly propose a solution based on data mining prediction to predict the best Triage category for the patients who utilize the ambulance services to reach the hospital at the time of an emergency situation.

5.4.2 Sub Research Question 1

It is essential to discover the real story behind the broad categories related to the relationship between the attributes of the pre-hospital ambulance service dataset to describe the real facts.

5.4.3 Sub Research Question 2

In Sri Lanka, there is not enough research that accurately analyzes the huge volume of pre-hospital data to inform decision-makers and policy-makers in the Ministry of Health, Nutrition and Indigenous Medicine to improve existing pre-hospital care services to the nations.

5.5 Summary

This chapter provides details on the design of the research and the applicability of the methods. In addition, this chapter focuses on high-level design for research and how under-research problems are structured in research. The following section will discuss the details of the implementation according to this design.

CHAPTER-6

Implementation

6.1 Introduction

In chapter 5 top-level design of the solution has been described. This chapter describes the implementation of each research question regarding data analysis software, process, methods, etc. This chapter is about how the research is implemented.

6.2 Solution for Sub Research Question-1

The 1990 Ambulance Service in Sri Lanka's pioneering pre-hospital care emergency response ambulance service. For every 100,000 people, one ambulance has been assigned and stationed in key locations of the districts.

As the solution for the sub-research question-1, the sufficient data set has been generated similar to the original dataset accordingly, in order to avoid the ethical clearance from both sides of the Ministry of Health, Nutrition & Indigenous Medicine, and the University of Moratuwa. The data set has been processed systematically in Rapid Miner and produces reports based on the necessity of the research accordingly.

Depending on the actual situation in the first year, the control center received more than 322,000 calls and treated more than 32,000 cases. The major cases being identified as Cardio, Trauma, and Stroke[19]. For this research based on the original dataset, the following Table-2 illustrates the suitable major categories and grouped with the subcategories for easy analysis.

#	Major Categories	Sub Categories
1	Advanced Cardiac Response (ACR)	Abnormal Cardiovascular Conditions
		Neurological Disorders
2	Sepsis	Pain Management
		Diabetic Emergencies
		Allergies and Anaphylaxis

		Poisoning/Overdose Emergencies
		Behavioral & Mental Health Emergencies
		Environmental Emergencies
3	Special Patient Groups	Pregnancy and Pre-Delivery and Post-Delivery Emergencies
		Childbirth and Neonatal Resuscitation
		Gynecological Emergencies
4	Pediatrics	Pediatric Medical
		Pediatric Trauma
5	Shock	Soft Tissue and Musculoskeletal Injuries including Crush Injuries
		Head, Brain and Spinal Injuries
6	Burn Injuries	Burn Injuries

Table 2: Major categories with subcategories

6.2.1 Status of the Dataset

Among the multiple calls, the number of calls which were selected for this process only the ambulances was saved to the patients. According to the processed data, the average number of calls per /day is 131.8 for the month of January and February.

The average response time (TLoca_time) and the dispatch time(TDis_time)) are respectively 10 minutes and 12 minutes as described in Figure-6.2.1.1 below. The "Row No"1, 2, 3, represent the month of January, February, and March respectively.

Row No. ↑	count(Date)	average(TLoca_time)	average(TDis_time)	average(TTime Taken)
1	4087	9.930	12.147	22.078
2	3693	10.015	12.960	22.975
3	2220	10.086	12.833	22.919

Figure-6.2.1.1: The average Response time and Dispatch time.

6.2.2 Clustering the Dataset

Clustering is the process of transforming a group of abstract objects into classes of similar objects. A cluster of data objects can be treated as a group. In cluster analysis, we first partition the dataset into groups based on the similarity of the data, then we assign the labels to the groups and apply the algorithms to find out the facts among the main categories of ambulance service. We conducted the cluster analysis for the dataset using Rapid Miner.

The Figure-6.2.2.1 below describes the cluster model of the data set and Figure-6.2.2.2 shows the performance vector of the model for the k value of 4 which has the lowest average within centroid distance.

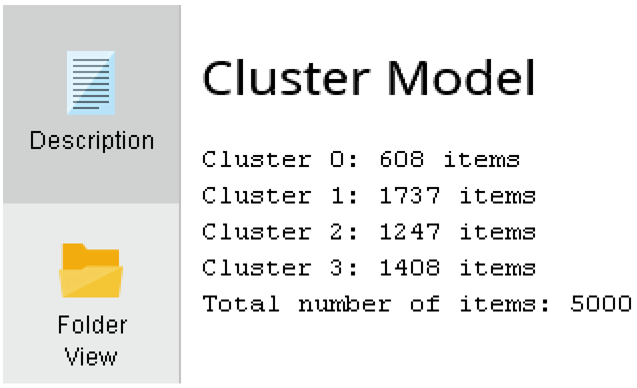


Figure-6.2.2.1: The cluster model.

The dataset was divided into 4 groups which Cluster-0 608 items, Cluster-1 1737 items, Cluster-2 1247 items, and Cluster-3 1408 items respectively. The Figure-6.2.2.2 is below describing each cluster model with the performance vector.

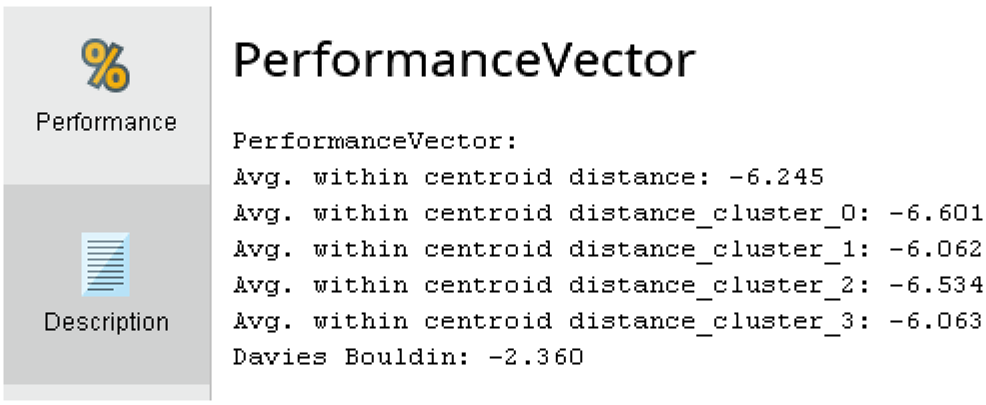


Figure-6.2.2.2: The performance vector of the cluster model

The below Figure-6.2.2.3 is describe the contents of each cluster model where the cluster-2 is the cluster for Critical requirement = Yes. The table provides multiple decisions based on the target role and the correlations among the variables.

Attribute	cluster_0	cluster_1	cluster_2 ↓	cluster_3
Critical_requirments = Yes	0.007	0.010	1	0.008
SEVERITY = Imminently Life Thretning	0	0	0.910	0
TRIAGE_CATEGORY = Orange	0	0	0.901	0
Sex = Female	0.536	0.626	0.512	0.448
Sex = Male	0.464	0.374	0.488	0.552
District = Gampaha	0.484	0.439	0.450	0.474
Incident_type = Stroke	0	0.021	0.306	0.022
District = Colombo	0.219	0.257	0.279	0.257
District = Kalutara	0.298	0.305	0.271	0.268
Incident_type = Heart disease	0	0.045	0.233	0.061
Incident_type = Accident	0.007	0.079	0.149	0.102
Incident_type = Chronic kidney disea	0.005	0.084	0.100	0.116
TRIAGE_CATEGORY = Red	0.007	0.010	0.099	0.008
SEVERITY = Life Threatning	0	0.001	0.081	0.001
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Figure-6.2.2.3: Contents of the cluster model

The Cluster_2 described that the data set having the critical condition for all 3 districts with a percentage of 45%, 27.9%, and 27.1% respectively in Gampaha, Colombo, and Kalutara. The incident types are available as Stroke (30.5%), Heart Diseases (23.3%) and Accident (14.9%), including both Male and Female patients.

6.2.3 Association for the Dataset

The association rule-extraction is a procedure for finding patterns, correlations, associations, or frequent causal structures from the datasets. For this research, we have conducted the association rule mining for the pre-hospital dataset to find out further evidence of the relationship among the attributes.

The Figure-6.2.3.1: below describing the association rule mining for the group of the patient in critical condition, triage category Red and Orange and the sex Male & Female.

Result History

AssociationRules (Create Association Rules)

Show rules matching

all of these conclusions:

CR_No
Tri_Yellow
S_Female
CR_Yes
Tri_Orange

No.	Premises	Conclusion	Support	Confid...	LaPlace	Gain	p-s	Lift	Conviction
55	Tri_Orange	CR_Yes	0.227	1	1	-0.227	0.171	4.036	∞
59	S_Female, Tri_Orange	CR_Yes	0.116	1	1	-0.116	0.087	4.036	∞
60	S_Male, Tri_Orange	CR_Yes	0.111	1	1	-0.111	0.083	4.036	∞
2	Tri_Orange	S_Female, CR_Yes	0.116	0.511	0.910	-0.338	0.087	4.011	1.785

Figure-6.2.3.1: Association Rule Mining for Patient Critical Condition-Yes

Furthermore, Figure 6.2.3.2: below represent the association rule mining for non-critical requirements as described below for the triage category of Green and Yellow. Therefore, it is highly concluded that the triage category of Orange and Red is the highly critical requirements patients carried by the ambulance staff to the Hospitals.

Show rules matching

all of these conclusions:

CR_No
Tri_Yellow
S_Female
CR_Yes
Tri_Orange

No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift	Convicti...
33	S_Male	CR_No	0.343	0.740	0.918	-0.584	-0.006	0.984	0.954
34	DT_Gampaha	CR_No	0.346	0.757	0.924	-0.568	0.002	1.006	1.018
35	S_Female, DT_Kalutara	CR_No	0.120	0.762	0.968	-0.194	0.002	1.013	1.040
36	S_Female	CR_No	0.409	0.763	0.917	-0.664	0.006	1.014	1.044
37	DT_Kalutara	CR_No	0.218	0.766	0.948	-0.352	0.004	1.018	1.058
38	S_Female, DT_Gampaha	CR_No	0.188	0.776	0.956	-0.297	0.006	1.032	1.108
50	Tri_Yellow , S_Female, DT_Gampaha	CR_No	0.156	0.999	1.000	-0.157	0.039	1.328	194.027
51	Tri_Yellow , DT_Gampaha	CR_No	0.285	0.999	1.000	-0.286	0.071	1.329	353.858
52	Tri_Yellow , S_Female	CR_No	0.343	0.999	1.000	-0.344	0.085	1.329	425.473
53	Tri_Yellow	CR_No	0.628	1.000	1.000	-0.629	0.156	1.329	778.835
54	Tri_Green	CR_No	0.124	1	1	-0.124	0.031	1.329	∞
56	Tri_Yellow , S_Male	CR_No	0.285	1	1	-0.285	0.071	1.329	∞

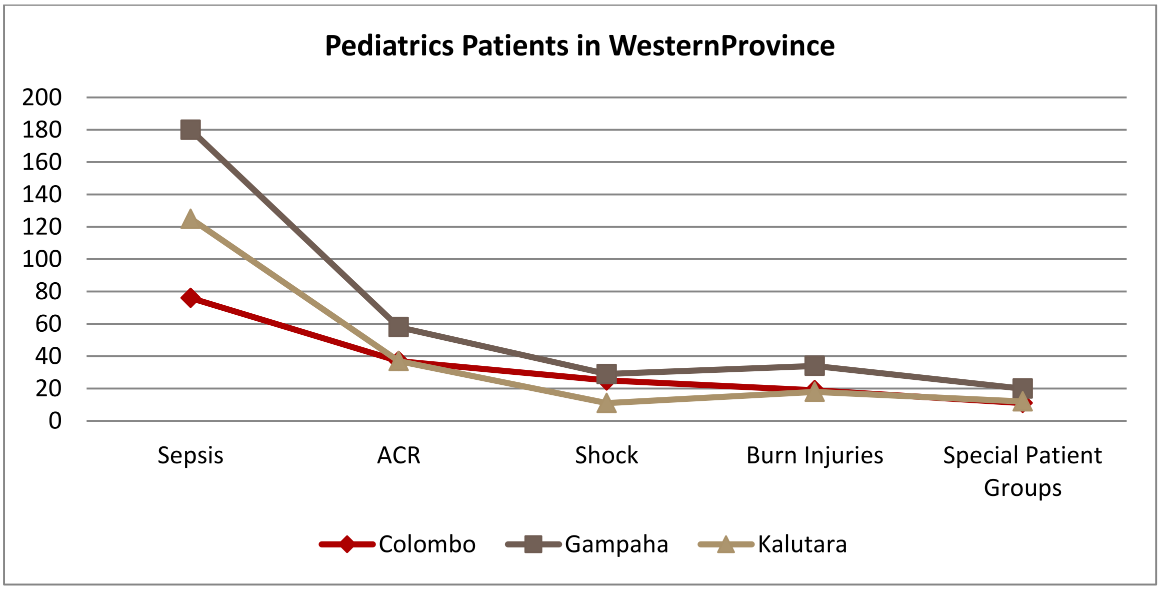
Figure-6.2.3.2: Association Rule Mining for Patient Critical Condition-No

Based on the above result, we can proceed to a further descriptive analysis as below, which can be used by Policymakers and or Decision makers to improve existing services.

6.2.4 Patient Group: Pediatrics (Age=<16)

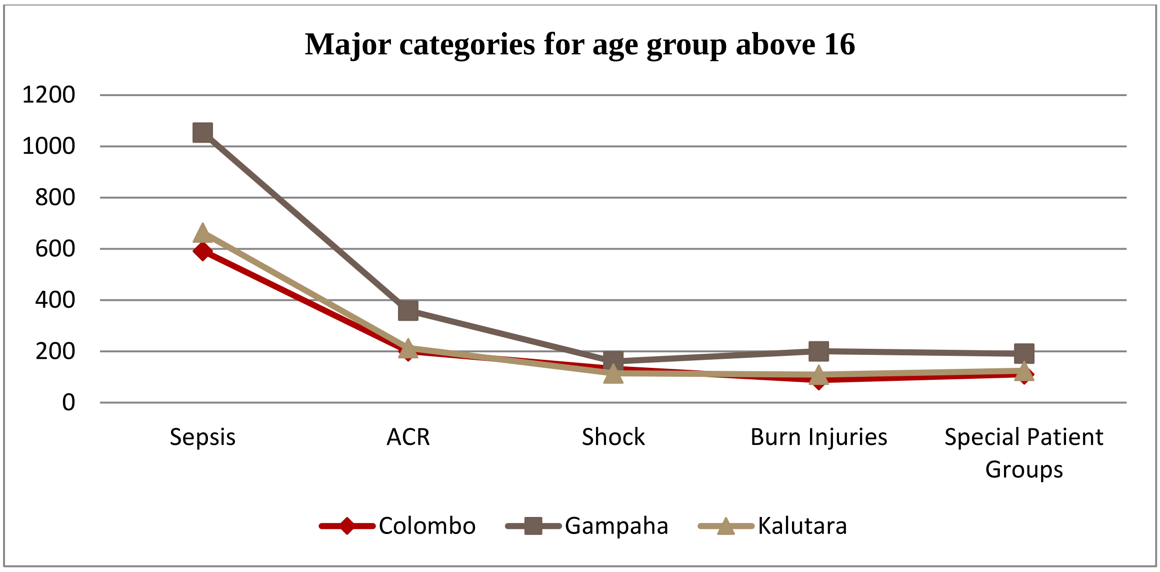
Pediatrics is the patient group that involves the Medical care of infants, children, and adolescents up to the age of 16 as a child patient. This patient is handled by the hospitals in a separate unit called "Pediatrics wards" for better care.

Pediatrics	Sepsis	ACR	Shock	Burn Injuries	Special Patient Groups
Colombo	76	37	25	19	11
Gampaha	180	58	29	34	20
Kalutara	125	37	11	18	12

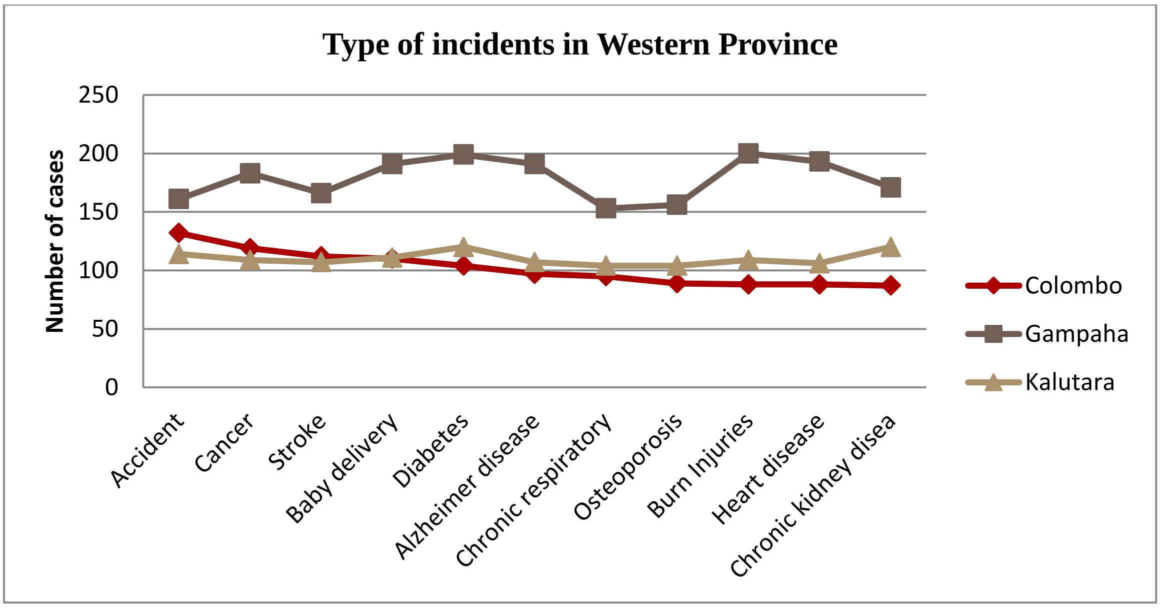


6.2.5 Patient Group: Age>16

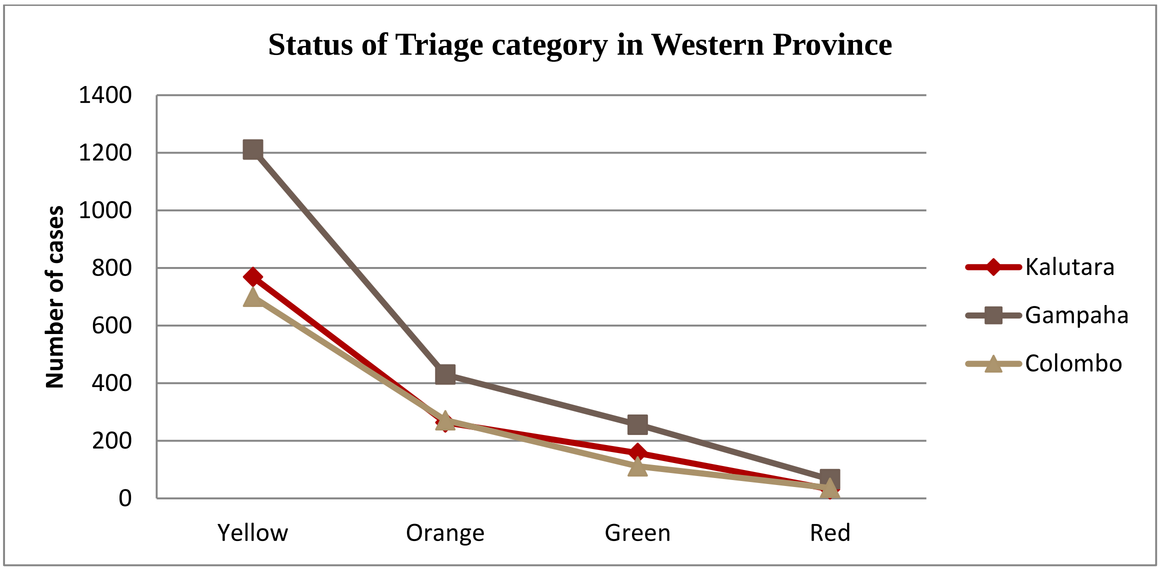
Age group above 16					
	Sepsis	ACR	Shock	Burn Injuries	Special Patient Groups
Colombo	591	200	132	88	110
Gampaha	1053	359	161	200	191
Kalutara	663	213	114	109	124



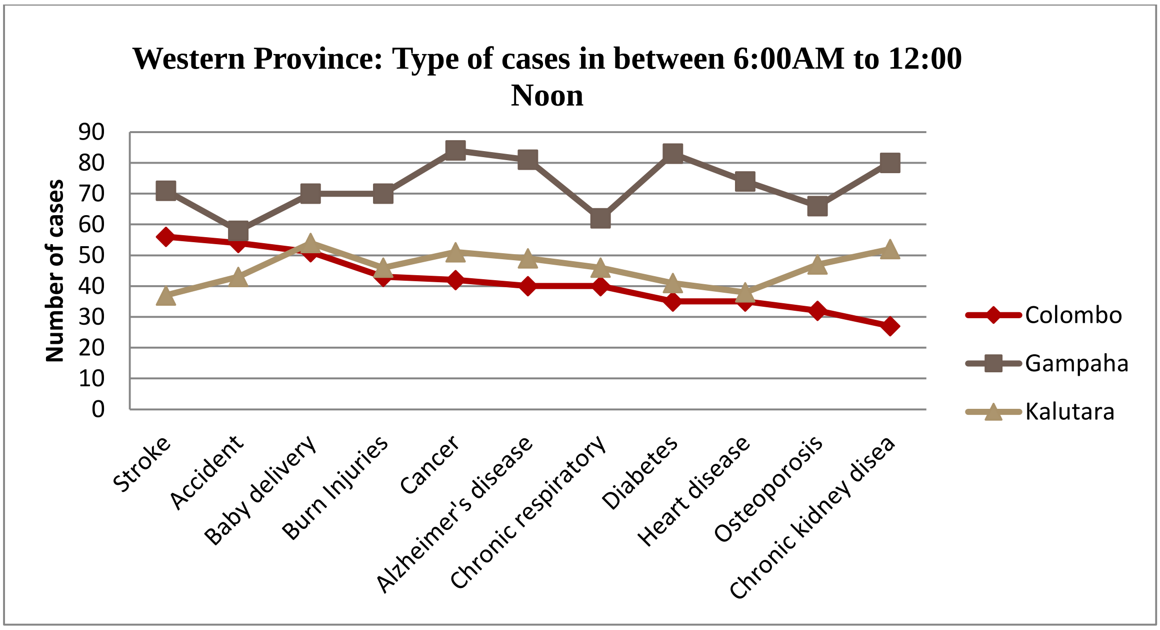
6.2.6 Type of Incidents in Western Province



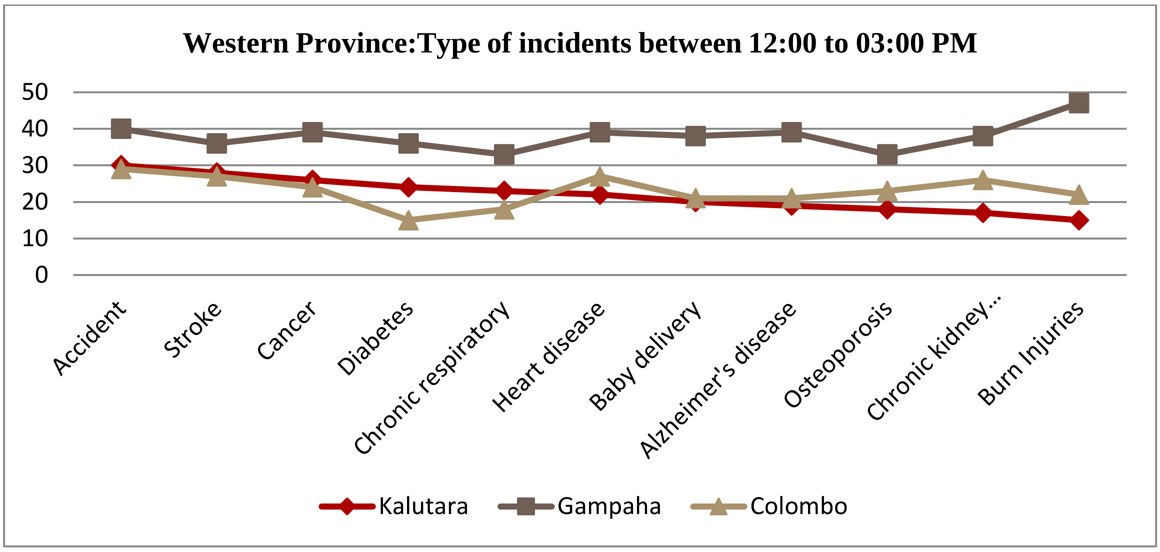
6.2.7 Triage category Status for the Western Province



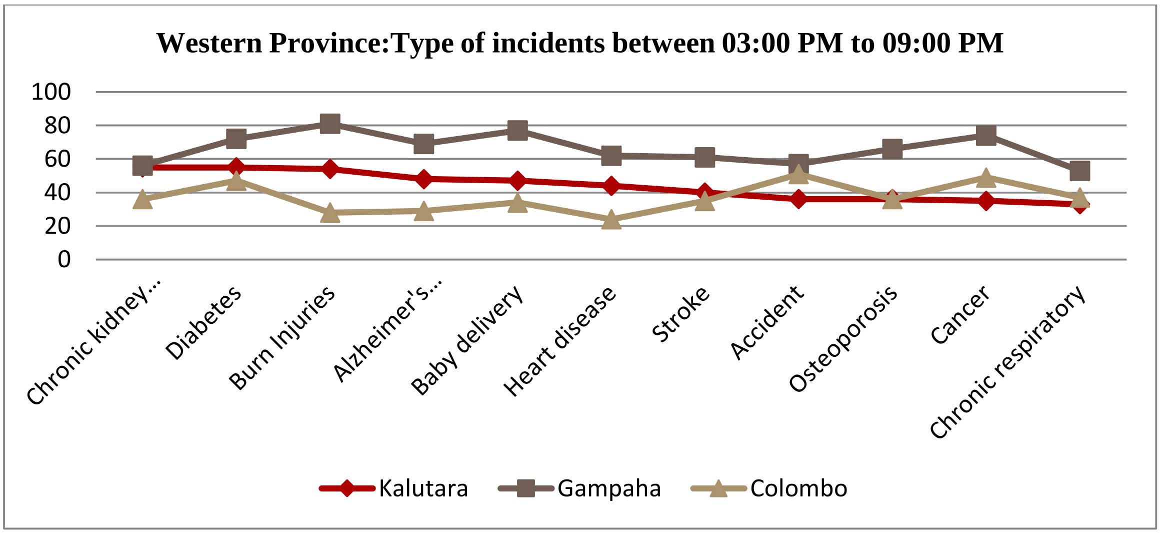
6.2.8 Type of cases between 06:00 AM to 12:00 Noon for the Western Province



6.2.9 Type of cases between 12:00 Noon to 03:00 PM for the Western Province



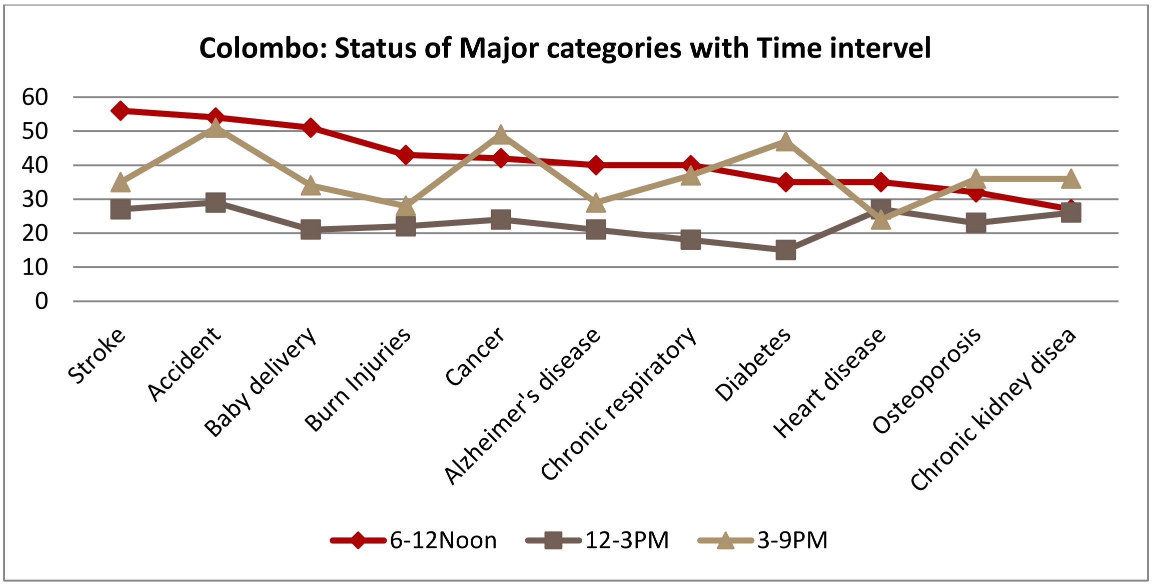
6.2.10 Type of cases between 03:00 PM to 09:00 PM for the Western Province



6.2.11 Major categories with Time interval in District wise

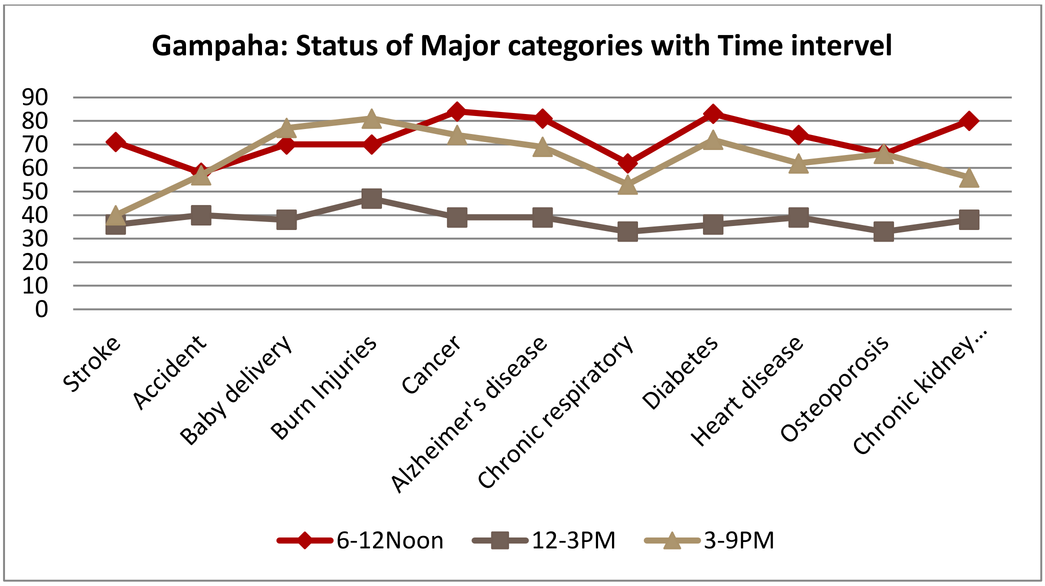
Status of Major categories with Time interval for Colombo Districts

Incident type	6-12Noon	12-3PM	3-9PM
Stroke	56	27	35
Accident	54	29	51
Baby delivery	51	21	34
Burn Injuries	43	22	28
Cancer	42	24	49
Alzheimer's disease	40	21	29
Chronic respiratory	40	18	37
Diabetes	35	15	47
Heart disease	35	27	24
Osteoporosis	32	23	36
Chronic kidney disease	27	26	36



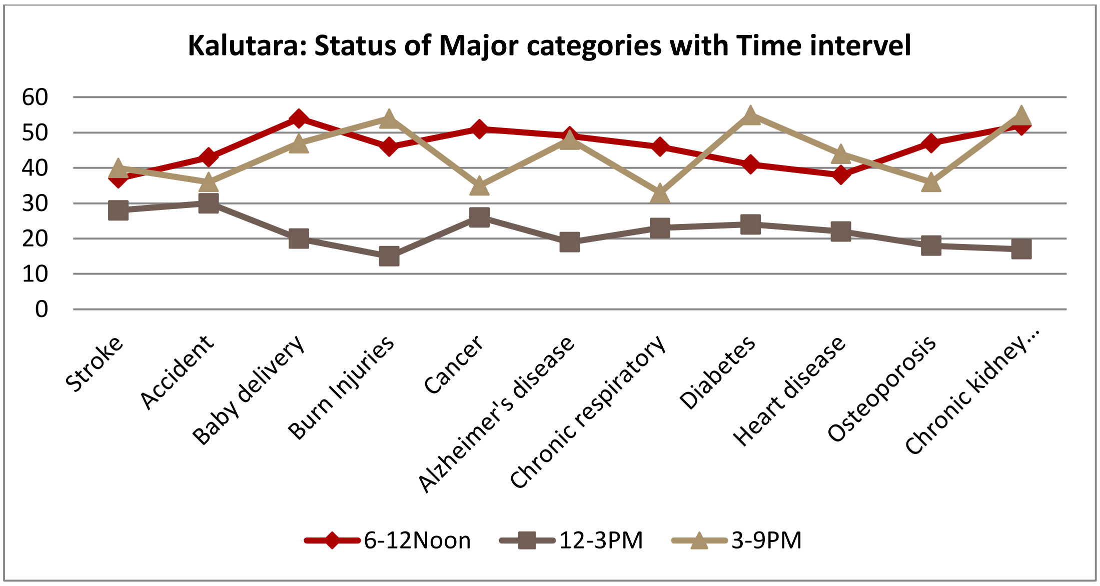
Status of Major categories with Time interval for Gampaha Districts

Incident type	6-12Noon	12-3PM	3-9PM
Stroke	71	36	40
Accident	58	40	57
Baby delivery	70	38	77
Burn Injuries	70	47	81
Cancer	84	39	74
Alzheimer's disease	81	39	69
Chronic respiratory	62	33	53
Diabetes	83	36	72
Heart disease	74	39	62
Osteoporosis	66	33	66
Chronic kidney disease	80	38	56



Status of Major categories with Time interval for Kalutara Districts

Incident type	6-12Noon	12-3PM	3-9PM
Stroke	37	28	40
Accident	43	30	36
Baby delivery	54	20	47
Burn Injuries	46	15	54
Cancer	51	26	35
Alzheimer's disease	49	19	48
Chronic respiratory	46	23	33
Diabetes	41	24	55
Heart disease	38	22	44
Osteoporosis	47	18	36
Chronic kidney disease	52	17	55



The detailed procedure is attached in Appendix-1 for reference based on the steps followed by the Rapid Miner.

6.3. Solution for Sub Research Question-2

Data or information on health care will not be enough to transform the results. Data, which are simple measures of patient characteristics and things, have little significance or inherent value. Data analysis allows the identification of models, thus creating information. Using the information to generate recommendations, action rules and behavior change means creating knowledge to make decisions and change human behavior.

Good decisions about effective policies, services and behaviors require up-to-date, accurate and relevant information. Health information is needed for strategic planning and priority setting to improve health services.

As the solution for the enhancement of the pre-health care services has to be analyzed in terms of past, present and future trends of the service in Sri Lanka. As per the studies of the current status, the ambulance services are updating the triage category by manually in their system which is inefficient for the actual figures of the triage category. The triage category depends on the acuity level of the time interval, where the patient has to be handover to the hospital by the ambulance according to the acuity range of timing interval for quality services.

The following technics will provide the predicted triage category based on the other factors correlated to the attributes within the dataset. The right prediction of the categories will enhance the ambulance services in the right direction.

As the solution for the sub research question of triage categories the classification method used to classify the dataset in order to predict the categories using the software Rapid Miner. In general, most researchers were used to perform the analysis using the KNN, Naïve Bayes and Decision Tree algorithms for analysis to accomplish the task [20].

6.3.1 KNN Classification

KNN is a non-parametric algorithm. The objective of this algorithm is to use a database in which the data points are separated into several classes to predict the classification of a new sample point. Hence, the algorithm was used for the ambulance data set to predict the classified values.

The Figure-6.3.1.1 below shows the KNN Classification provided the **accuracy of 91.44% +/- 0.47% (micro average: 91.44%)** by assigning the k value of 18.

ExampleSet (Set Role) | KNNClassification (k-NN) | PerformanceVector (Performance)

Table View | Plot View

accuracy: 91.44% +/- 0.47% (micro average: 91.44%)

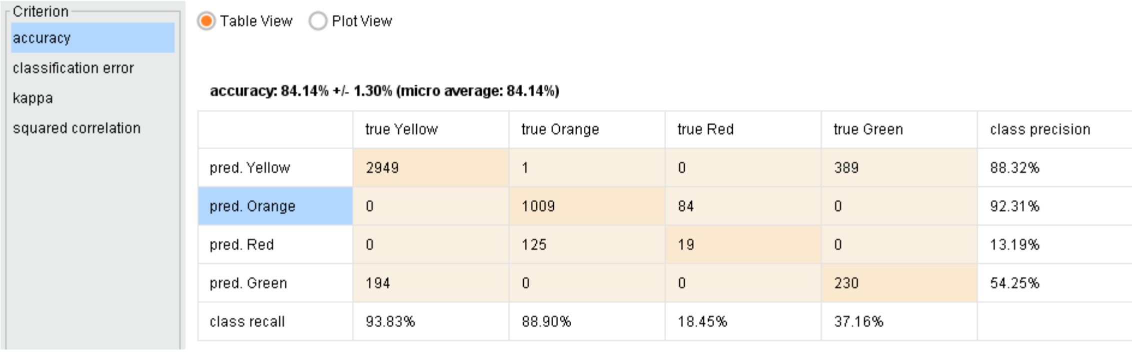
	true Yellow	true Orange	true Red	true Green	class precision
pred. Yellow	3089	57	5	207	91.99%
pred. Orange	4	1071	98	0	91.30%
pred. Red	0	1	0	0	0.00%
pred. Green	50	6	0	412	88.03%
class recall	98.28%	94.36%	0.00%	66.56%	

Figure -6.3.1.1: The performance vector of the KNN cluster model

6.3.2 Naïve-Bayes Classification

A Naïve-Bayes classifier is a probabilistic model of machine learning used for the classification task. The classifier's node is based on Bayes' theorem.

The Figure-6.3.2.1 below describes the Naïve Bayes classification provided the **accuracy of 84.14% +/- 1.30% (micro average: 84.14%)**



accuracy: 84.14% +/- 1.30% (micro average: 84.14%)

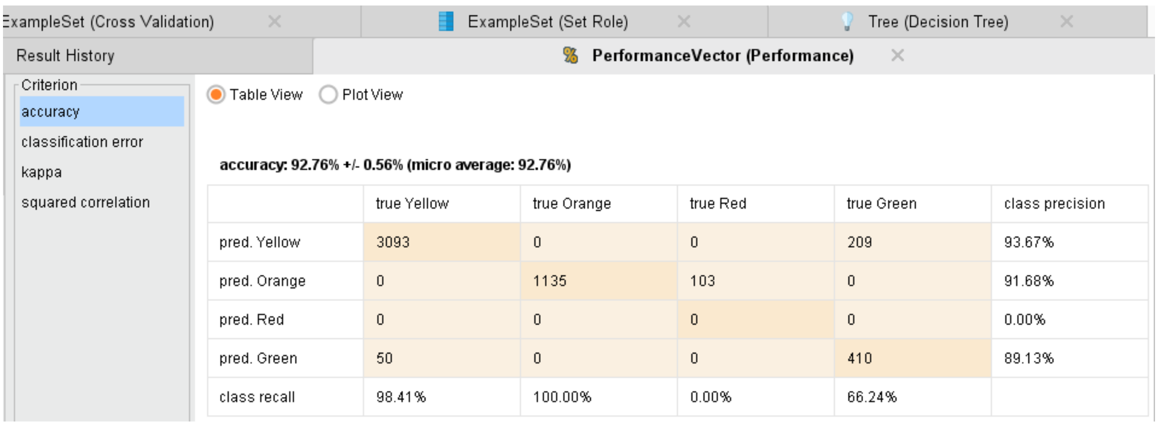
	true Yellow	true Orange	true Red	true Green	class precision
pred. Yellow	2949	1	0	389	88.32%
pred. Orange	0	1009	84	0	92.31%
pred. Red	0	125	19	0	13.19%
pred. Green	194	0	0	230	54.25%
class recall	93.83%	88.90%	18.45%	37.16%	

Figure-6.3.2.1: The performance vector of the Naive Bayes cluster model

6.3.3 Decision Tree Classification

In computer science, a decision tree (Predictive model) for use in data mining, a decision tree is used to learn a classification function that predicts the value of a dependent attribute based on the values of independent variables where the classification tree can be an input for decision making.

The following Figure-6.3.3.1 describes the performance vector of the cluster model. The performance of the Tree model has predicted the best triage category for the individuals of the dataset with the **accuracy of 92.76% +/- 0.56% (micro average: 92.76%)**



accuracy: 92.76% +/- 0.56% (micro average: 92.76%)

	true Yellow	true Orange	true Red	true Green	class precision
pred. Yellow	3093	0	0	209	93.67%
pred. Orange	0	1135	103	0	91.68%
pred. Red	0	0	0	0	0.00%
pred. Green	50	0	0	410	89.13%
class recall	98.41%	100.00%	0.00%	66.24%	

Figure-6.3.3.1: The performance vector of Decision Tree cluster model

Multiple models have been tested accordingly among them the best classification method of higher accuracy of the model (92.76%) has been identified as the decision tree model. The below Figure-6.5.3.2: describe the sample of predictive values for the triage category for reference.

Row No.	TRIAGE CAT...	prediction(TRIAGE CATEGORY)	confidence(Red)	confidence(Green)	confidence(Orange)	confidence(Yellow)
1	Red	Orange	0.000	0.000	1.000	0.000
2	Yellow	Yellow	0.000	0.070	0.000	0.930
3	Green	Yellow	0.000	0.042	0.000	0.958
4	Green	Green	0.000	1.000	0.000	0.000
5	Red	Orange	0.002	0.000	0.998	0.000
6	Yellow	Green	0.000	0.992	0.000	0.008
7	Yellow	Yellow	0	0.160	0.000	0.840
8	Green	Yellow	0.000	0.042	0.000	0.958
9	Yellow	Yellow	0.000	0.000	0.000	1.000
10	Yellow	Yellow	0.000	0.000	0.000	1.000
11	Orange	Orange	0.000	0.000	1.000	0.000
12	Orange	Orange	0.492	0.005	0.499	0.004
13	Orange	Red	0.972	0.000	0.028	0.000

ExampleSet (5,000 examples, 6 special attributes, 9 regular attributes)

Figure-6.3.3.2: The predicted triage category

The detailed procedures are attached in Appendix-2 for reference based on the procedures followed by the Rapid Miner.

6.3.4 Incident locations in frequency based

Currently, there were 56 ambulances located in different locations of the Western Province to provide the emergency ambulance services to the public. The movements of the ambulance were monitored by the GPS technology for logging, controlling and advising the traffic route to the driver for avoiding the delay routes when they carry the patient to the hospital.

The current problem is that ambulance locations have been assigned taking into account security locations but not targeted locations. By improving these locations, security locations will become targeted locations, which will also improve response time and dispatch time as soon as possible.

As the solutions for the sub research question the data set has been processed and found the top most frequent ambulance requested area in each district separately, where the incident location with the frequent occurrence is listed in the Figure-6.4.1: below as describe the topmost incident locations for Colombo District.

District	Incient_Loc...	Incident_type	Triage Cate...	Row No	Frequency ↓
Colombo	Bambalapitiya	Accident	Orange	35	5
Colombo	Bopetta	Heart disease	Orange	311	5
Colombo	Udagama	Stroke	Orange	72	5
Colombo	Avissawella	Stroke	Orange	104	4
Colombo	Boralesgamu...	Osteoporosis	Orange	51	4
Colombo	Mawittara Nor...	Stroke	Orange	61	4
Colombo	Nawala East	Stroke	Orange	198	4
Colombo	Dehiwala West	Stroke	Orange	101	3

Figure-6.4.1: The topmost frequent places in Colombo District

Based on the data analysis Bambalapitiya, Bopetta and Udagama having 5 in frequency and Avissawella, Borlesgamuwa, Mawittara North, and Nawala East having 4 in frequency in Colombo District.

The Figure-6.4.2 below shows the topmost frequent places in Kalutara Districts where the topmost incident location is Govinna-North having 6 in frequency for the occurrence.

..	District	Incient_Loc...	Incident_type	Triage Cate...	Row No	Frequency ↓
.	Kalutara	Govinna North	Heart disease	Orange	158	6
.	Kalutara	Aluthgama	Heart disease	Red	134	4
.	Kalutara	Katukurundu...	Heart disease	Orange	103	4
.	Kalutara	Kuda Payaga...	Heart disease	Orange	194	4
.	Kalutara	Mathugama ...	Heart disease	Orange	131	4
.	Kalutara	Molligoda So...	Stroke	Orange	237	4
.	Kalutara	Pokunuwita	Accident	Orange	207	4
.	Kalutara	Punchideniya	Accident	Red	3	4

Figure-6.4.2: The topmost frequent places in Kalutara District

The Figure-6.4.3 below shows the topmost frequent places in Gampaha District where the Bothale-Pahalagama, Gonahena-East, and Opathella are having 6 in frequency occurrence for the triage category of Orange. The decision-makers have to be special attention in this area for future consideration.

..	District	Incient_Location	Incident_type	Triage Cate...	Row No	Frequency ↓
.	Gampaha	Bothale Pahalagama	Cancer	Orange	59	6
.	Gampaha	Gonahena East	Heart disease	Orange	19	6
.	Gampaha	Opathella	Stroke	Orange	141	6
.	Gampaha	Dalupitiya East	Stroke	Orange	131	5
.	Gampaha	Eldeniya East	Stroke	Orange	171	5
.	Gampaha	Kadawala North	Heart disease	Orange	295	5
.	Gampaha	Neligama	Stroke	Orange	351	5
.	Gampaha	Polwatta	Heart disease	Orange	204	5
.	Gampaha	Thalahena	Accident	Orange	47	5
.	Gampaha	Batuwatta East	Chronic respiratory	Orange	299	4
.	Gampaha	Divilupitiya	Cancer	Orange	2	4
.	Gampaha	Galborella	Chronic respiratory	Orange	167	4
.	Gampaha	Ganepola	Stroke	Orange	112	4
.	Gampaha	Induragara North	Heart disease	Orange	36	4

Figure-6.4.3: The topmost frequent places in Gampaha District

6.5 Summary

The Implementation chapter provides the complete path in constructing data models to handle the research sub-question. In addition, this chapter gives a detailed description of the answers provided. The next chapter will deal with the evaluation of the algorithms used in this chapter.

CHAPTER-7

Evaluation

7.1 Introduction

This chapter discusses the test strategies used for the research question on measurement metrics for selected data mining techniques.

7.2 Evaluation for Classification

The confusion matrix in Figure -7.2.1 shows the number of correct and incorrect predictions made by the classification model against the actual results (target value) in the data. The matrix is NxN, where N is the number of target values (classes). The performance of such models is generally evaluated using the matrix data.

Confusion Matrix		Target			
		Positive	Negative		
Model	Positive	a	b	*Positive Predictive Value*	a/(a+b)
	Negative	c	d	*Negative Predictive Value*	d/(c+d)
		Sensitivity	*Specificity*	**Accuracy** = (a+d)/(a+b+c+d)	

Figure -7.2.1: The Detail of the Matrix

- Accuracy: the proportion of the total number of correct predictions.

- Positive predictive value or precision: the proportion of positive cases correctly identified.

- Negative predictive value: the proportion of correctly identified negative cases.

- Sensitivity or recall: the proportion of actual positive cases correctly identified.

- Specificity: the proportion of actual negative cases correctly identified.

7.2 Evaluation for Association

Association Rules Mining (ARM) has attracted a lot of attention over the last decade. It aims to extract a set of relevant rules from a given database. In order to evaluate the quality of the resulting rules, existing measures, such as support and trust/confidence, allow the rules resulting from the ARM process to be evaluated separately, by omitting the different dependencies between the rules.

Most of the researchers were used to evaluate the association rule mining based on the execution time of the dataset. The algorithms most commonly used for this type of action are Apriori and FP-Growth. Performance analysis is performed based on the number of instances and the level of trust. The effectiveness of the two algorithms is evaluated according to the time needed to generate the association rules. The experimental data presented allow us to conclude that the FP-Growth algorithm behaves better than the Apriori algorithm [21].

7.3 Evaluation for Clustering

To assess the correctness of the data model the data set processed in Rapid Miner tool and then K-means clustering algorithm is used to evaluate the data set with classes to cluster evaluation option. The following Figure-7.3.1 describe the sample output of a cluster model where the value of k=4 (The number of clusters =4).

id	Major Categ...	cluster	Sex = Female	Sex = Male	District = Co...	District = Ka...	District = Ga...	Incident_loc
1	Sepsis	cluster_1	1	0	1	0	0	1
2	ACR	cluster_2	0	1	1	0	0	0
3	Sepsis	cluster_2	1	0	1	0	0	0
4	Sepsis	cluster_3	0	1	1	0	0	0
5	ACR	cluster_2	1	0	0	1	0	0
6	ACR	cluster_2	0	1	1	0	0	0
7	Sepsis	cluster_0	1	0	0	0	1	0
8	Sepsis	cluster_3	1	0	0	0	1	0
9	Sepsis	cluster_0	1	0	0	1	0	0
10	Sepsis	cluster_2	0	1	1	0	0	0
11	Sepsis	cluster_0	0	1	0	1	0	0
12	Speial Patien...	cluster_1	1	0	0	0	1	0
13	ACR	cluster_2	0	1	0	1	0	0

Figure -7.3.1: The output of a cluster model where the k=4

To decide the number of k values or cluster values multiple k values applied to the data set to generate the average within centroid distance measurements as in Table 2 below.

K Value	2	3	4	5	6	7	8
Avg. within centroid distance	7.266	6.73	6.245	6.015	5.827	5.705	5.591

Table 3: K-Values Vs Average Centroid Distance

The Elbow method, used by researchers, gives an idea of how many k clusters would be based on the sum of the squared distance between data points and centroids in their cluster [22].

We choose k at the point where (k=4) the mean in the centroid distance begins to flatten and form an elbow as in Figure:7.2 below illustrates the flow of the K values Vs average centroid distance.

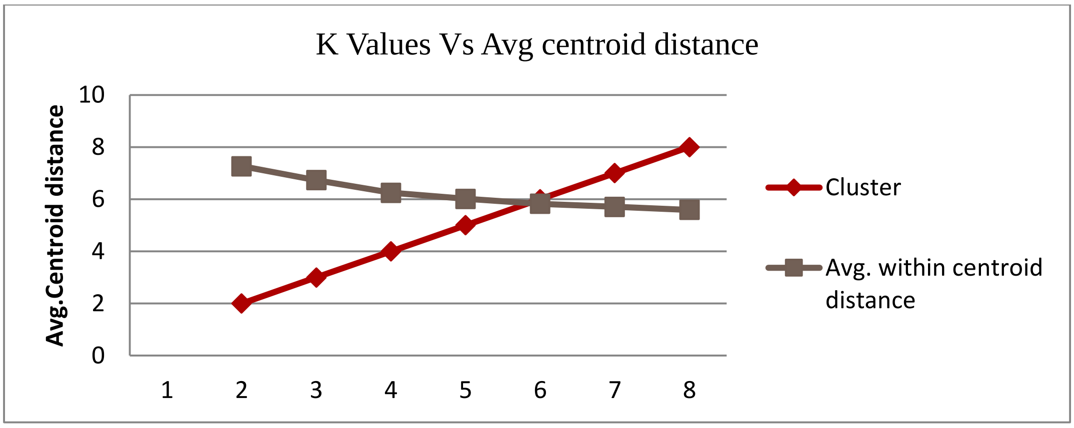


Figure 7.2: K-Values Vs Average Centroid Distance

7.4 Summary

This chapter ends with the results used to evaluate the data model. The final chapter will summarize all the research and highlight the main results of the research.

CHAPTER-8

Conclusion and Further Work

8.1 Introduction

This chapter delivers an overview of the research and how we provide the solution to address the problem of analyzing pre-hospital care ambulance service. Furthermore, this chapter focusses of limitations and further work of this research.

8.2 Overview of the research

By analyzing the ambulance service for pre-hospital care in Sri Lanka, we can discover the problems related to data that is not available in public. Patient data should be treated with ethical permission from the health directorate, which is very difficult to obtain in time for research. Due to these issues, the required data was generated using tools and analyzed accordingly to fully fulfill the purpose of this research.

The aim of this research is to improve existing services by analyzing the pre-hospital dataset. Data mining is identified as the best approach for discovering hidden models in this pre-hospital dataset. Different data mining procedures have followed to resolve the problem of under-research. The main paradigms of data mining are predictive techniques and descriptive techniques depending on the desired result. To determine the accuracy of the solution, different algorithms are used in the selected techniques and their comparative efficiency before selecting the best approach to draw the conclusions.

The Health Ministry should be considered to provide the original dataset for this type of Big Data analysis. We can apply the same techniques to analyze whether the original dataset provided by the Ministry of Health, Nutrition and Indigenous Medicine. Generated reports and decisions will motivate decision-makers to improve services in the right direction.

8.3 Problem encountered & limitations

The main objective of this research is to analyze data from the ambulance service for prehospital care generated by a computer tool. Where the data generated is slightly different from the original. As a result, the number of search problems that we can focus on is limited by the attributes of the dataset.

When we analyze the dataset we have encountered the following Problems / Limitations which has to be addressed by the Ministry of Health Decision-makers & Policymakers.

- There is no proper communication channel between the ambulance team and the hospital management to update the patient's status. Hospital arrival status for advance preparation should be arranged by hospital staff.

- Most of the time the ambulance team doesn’t know what is the status of the patient delivered to the hospital due to lack of communication between Ambulance Management team and the Hospital Administration in order to update the patient status from the ambulance management side.

- The state sector hospitals are running with the manual systems and don't have an either Hospital Management System or Patient Management System in practice which leads to the barrios of all the communications.

- The current ambulance stations have to be changed from safety location to target-oriented places will also enhance the response time in the future.

8.4 Further work

This research only addresses the research questions related to findings of the status and targeting to enhance the pre-hospital care ambulance service in Western Province of Sri Lanka. The original data may contain some other attributes which will provide hidden patterns for enhancing and developing or predicting the new items based on the dataset.

Hence, this research can be extended further to address different issues in pre-hospital care ambulance services in Sri Lanka.

8.5 Summary

This chapter concludes that describing the solution given with data mining to analyze the existing pre-hospital care system and how it can be enhanced further to improve the level of accuracy in predicting /exploring the pre-hospital data.

Reference

[1] I. Medicine and S. Lanka, "GUIDELINES FOR ACCIDENT AND EMERGENCY CARE SERVICES IN GOVERNMENT HOSPITALS," 2016.

[2] R. Fairhurst, "Pre-hospital care in Europe.," *Emerg. Med. J.*, vol. 22, no. 11, p. 760, 2005.

[3] E. Pitt and A. Pusponegoro, "Prehospital care in Indonesia.," *Emerg. Med. J.*, vol. 22, no. 2, pp. 144–7, 2005.

[4] M. Boyle and V. Plummer, "The pre-hospital and healthcare system in Malang , Indonesia Review The pre-hospital and healthcare system in Malang , Indonesia," vol. 14, no. 2.

[5] A. Sarlan, F. K. Xiong, R. Ahmad, W. Fatimah, W. Ahmad, and E. Bhattacharyya, "Pre-hospital Emergency Notification System," vol. 2015, pp. 168–173, 2015.

[6] J. A. Razzak and A. L. Kellermann, "Emergency medical care in developing countries: Is it worthwhile?," *Bull. World Health Organ.*, vol. 80, no. 11, pp. 900–905, 2002.

[7] Ministry of Health - Sri Lanka, "Accident and Emergency care Policy of Sri Lanka," 2016.

[8] Ministry of Health, "National policy and strategic framework on injury prevention & management in Sri Lanka," 2010.

[9] K. Wimalaratne, J. IL Lee, K. H. Lee, H. Y. Lee, J. H. Lee, and I. H. Kang, "Emergency medical service systems in Sri Lanka: problems of the past, challenges of the future," *Int. J. Emerg. Med.*, vol. 10, no. 1, p. 10, 2017.

[10] A. Akram, M. Anjum, M. Rehman, H. Bukhary, H. Amir, and R. Qaisar, "Life Savior : An Integrated Emergency Response System," pp. 1002–1006, 2017.

[11] A. Ahmed, A. Ishaque, and T. Nawaz, "Information and Communication Technology Introducing Efficiency in Emergency Medical Services," *2014 IEEE Int. Conf. Manag. Innov. Technol.*, pp. 211–215, Sep. 2014.

[12] P. Care, "Evaluation of emergency medical services systems: a classification to assist in the determination of indicators," pp. 1997–2000, 2003.

[13] S. Dharmaratne, A. Jayatilleke, and A. Jayatilleke, "WHO | Road traffic crashes, injury and fatality trends in Sri Lanka: 1938–2013," *Bull. World Health Organ.*, vol. 93, no. April, pp. 640–647, 2015.

[14] K. Williamson, R. Ramesh, and A. Grabinsky, "Advances in prehospital trauma care," *Int J Crit Illn Inj Sci*, vol. 1, no. 1, pp. 44–50, 2011.

[15] J. S. Davis *et al.*, "An analysis of prehospital deaths," *J. Trauma Acute Care Surg.*, vol. 77, no. 2, pp. 213–218, 2014.

[16] M. Bigdeli, D. Khorasani-Zavareh, and R. Mohammadi, "Pre-hospital care time intervals among victims of road traffic injuries in Iran_A cross-sectional study," *BMC Public Health*, vol. 10, 2010.

[17] C. W. Seymour, T. D. Rea, J. M. Kahn, A. J. Walkey, D. M. Yealy, and D. C. Angus, "Severe Sepsis in Pre-Hospital Emergency Care Analysis of Incidence, Care, and Outcome."

[18] S. Trauma, "' Big data ' approaches to trauma outcome prediction and autonomous resuscitation," vol. 75, no. 11, pp. 637–641, 2014.

[19] www.ft.lk, "1990 Suwasariya' Ambulance Service celebrates the first anniversary," *Monday, 31 July 2017 00:08*, 2017. [Online]. Available: http://www.ft.lk/healthcare/1990-suwasariya-ambulance-service-celebrates-first-anniversary/45-632293. [Accessed: 08-Jul-2019].

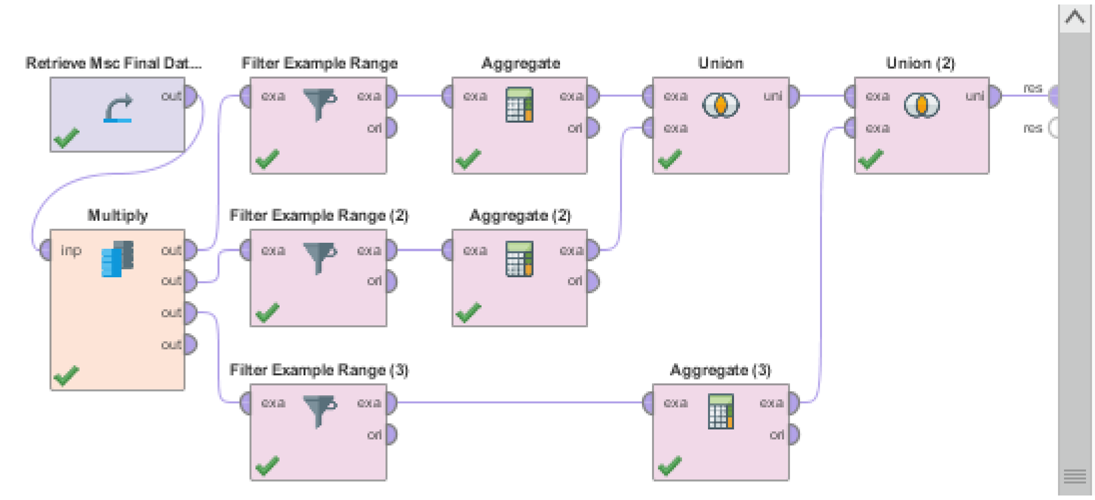
[20] C. S. Dangare and M. E. Cse, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques," vol. 47, no. 10, pp. 44–48, 2012.

[21] I. K.Vanitha, R.Santhi Department of Computer Studies, Saranathan College of Engineering, Trichy, "EVALUATING THE PERFORMANCE OF ASSOCIATION RULE MINING," vol. 2, no. 6, pp. 101–103, 2011.

[22] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, "Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 336, no. 1, 2018.

Appendixture-1

6.2.1 Status of the Dataset

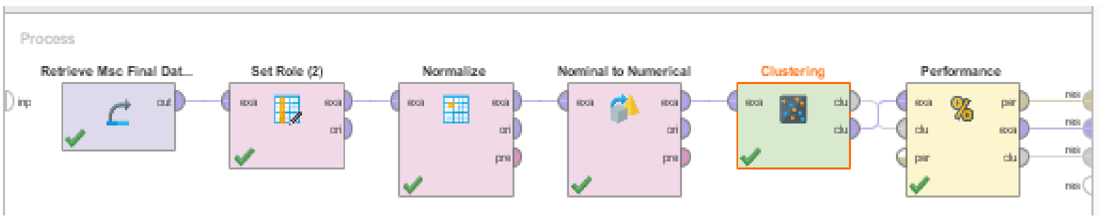


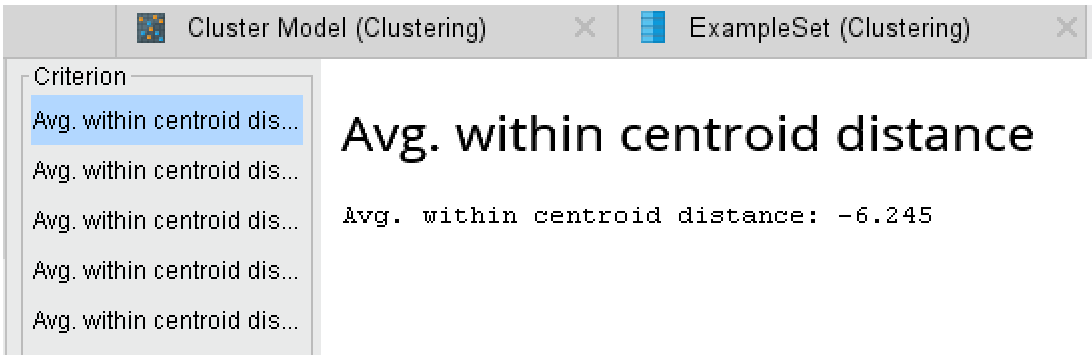
The output of the above process

Open in Turbo Prep Auto Model

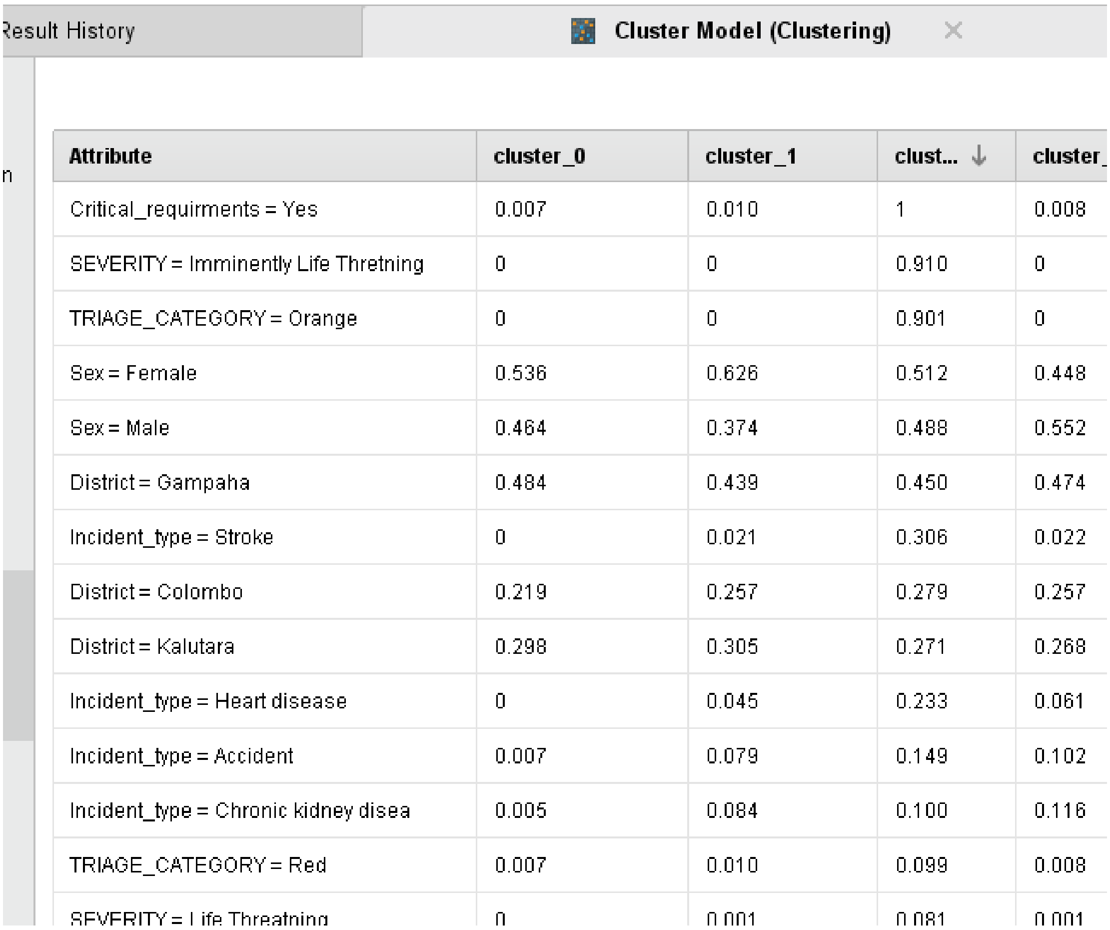
Row No.	count(Date)	average(TLo...	average(TDi...	average(TTi...
1	4087	9.930	12.147	22.078
2	3693	10.015	12.960	22.975
3	2220	10.086	12.833	22.919

6.2.2 Clustering the Dataset



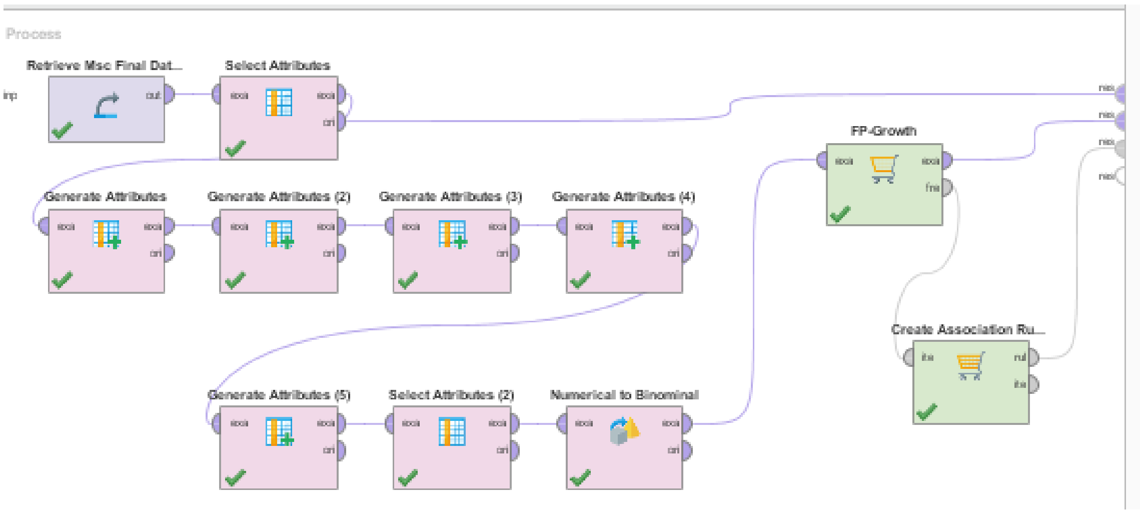


The cluster model for the dataset



Attribute	cluster_0	cluster_1	clust...	cluster_
Critical_requirments = Yes	0.007	0.010	1	0.008
SEVERITY = Imminently Life Thretning	0	0	0.910	0
TRIAGE_CATEGORY = Orange	0	0	0.901	0
Sex = Female	0.536	0.626	0.512	0.448
Sex = Male	0.464	0.374	0.488	0.552
District = Gampaha	0.484	0.439	0.450	0.474
Incident_type = Stroke	0	0.021	0.306	0.022
District = Colombo	0.219	0.257	0.279	0.257
District = Kalutara	0.298	0.305	0.271	0.268
Incident_type = Heart disease	0	0.045	0.233	0.061
Incident_type = Accident	0.007	0.079	0.149	0.102
Incident_type = Chronic kidney disea	0.005	0.084	0.100	0.116
TRIAGE_CATEGORY = Red	0.007	0.010	0.099	0.008
SEVERITY = Life Threatning	0	0.001	0.081	0.001

6.2.12 Association for the Dataset



The output of the association

Result History

AssociationRules (Create Association Rules)

Show rules matching

all of these conclusions:

CR_No
Tri_Yellow
S_Female
CR_Yes
Tri_Orange

Min. Criterion:

No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift	Convicti...
17	S_Male, DT_Gampaha	Tri_Yellow	0.129	0.601	0.930	-0.300	-0.006	0.956	0.931
18	S_Male, DT_Gampaha	CR_No , Tri_Yellow	0.129	0.601	0.930	-0.300	-0.006	0.957	0.932
19	S_Male	Tri_Yellow	0.285	0.615	0.878	-0.642	-0.006	0.979	0.966
20	S_Male	CR_No , Tri_Yellow	0.285	0.615	0.878	-0.642	-0.006	0.979	0.966
21	DT_Gampaha	CR_No , Tri_Yellow	0.285	0.625	0.882	-0.629	-0.002	0.994	0.990
22	DT_Gampaha	Tri_Yellow	0.286	0.625	0.882	-0.628	-0.002	0.994	0.990
23	DT_Colombo	Tri_Yellow	0.161	0.625	0.923	-0.354	-0.001	0.995	0.991
24	DT_Colombo	CR_No , Tri_Yellow	0.161	0.625	0.923	-0.354	-0.001	0.995	0.992
25	DT_Kalutara	Tri_Yellow	0.182	0.637	0.920	-0.389	0.003	1.014	1.024
26	DT_Kalutara	CR_No , Tri_Yellow	0.182	0.637	0.920	-0.389	0.003	1.014	1.025
27	S_Female	CR_No , Tri_Yellow	0.343	0.640	0.874	-0.730	0.006	1.018	1.031
28	S_Female	Tri_Yellow	0.343	0.640	0.874	-0.730	0.006	1.018	1.032
29	S_Female, DT_Gampaha	CR_No , Tri_Yellow	0.156	0.645	0.931	-0.328	0.004	1.027	1.047

The output of the association for Critical condition yes

Show rules matching

all of these conclusions:

CR_No
Tri_Yellow
S_Female
CR_Yes
Tri_Orange

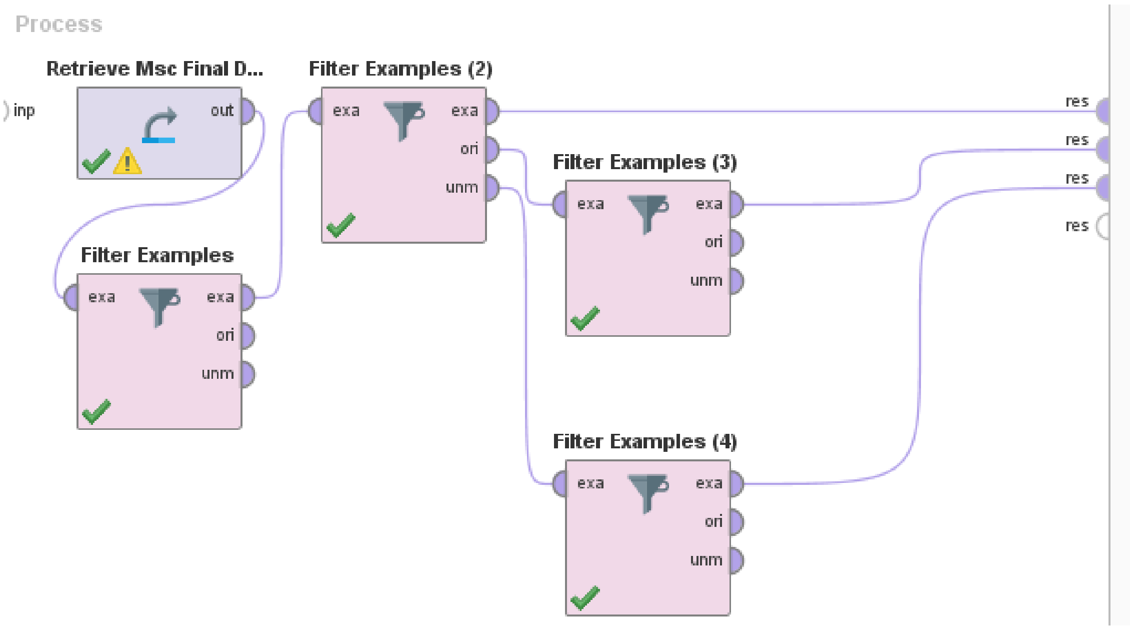
No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s
55	Tri_Orange	CR_Yes	0.227	1	1	-0.227	0.171
59	S_Female, Tri_Orange	CR_Yes	0.116	1	1	-0.116	0.087
60	S_Male, Tri_Orange	CR_Yes	0.111	1	1	-0.111	0.083

The association rules for the dataset

```
[Tri_Orange ] --> [S_Female] (confidence: 0.511)
[Tri_Orange ] --> [S_Female, CR_Yes ] (confidence: 0.511)
[CR_Yes , Tri_Orange ] --> [S_Female] (confidence: 0.511)
[CR_Yes ] --> [S_Female] (confidence: 0.514)
[DT_Gampaha] --> [S_Female] (confidence: 0.530)
[DT_Colombo] --> [S_Female] (confidence: 0.532)
[CR_No ] --> [S_Female] (confidence: 0.544)
[CR_No , DT_Gampaha] --> [S_Female] (confidence: 0.544)
[Tri_Yellow ] --> [CR_No , S_Female] (confidence: 0.546)
[CR_No , Tri_Yellow ] --> [S_Female] (confidence: 0.546)
[Tri_Yellow ] --> [S_Female] (confidence: 0.546)
[CR_No , DT_Kalutara] --> [S_Female] (confidence: 0.548)
[Tri_Yellow , DT_Gampaha] --> [CR_No , S_Female] (confidence: 0.548)
[CR_No , Tri_Yellow , DT_Gampaha] --> [S_Female] (confidence: 0.548)
[Tri_Yellow , DT_Gampaha] --> [S_Female] (confidence: 0.548)
[DT_Kalutara] --> [S_Female] (confidence: 0.550)
[S_Male, DT_Gampaha] --> [Tri_Yellow ] (confidence: 0.601)
[S_Male, DT_Gampaha] --> [CR_No , Tri_Yellow ] (confidence: 0.601)
[S_Male] --> [Tri_Yellow ] (confidence: 0.615)
[S_Male] --> [CR_No , Tri_Yellow ] (confidence: 0.615)
[DT_Gampaha] --> [CR_No , Tri_Yellow ] (confidence: 0.625)
[DT_Gampaha] --> [Tri_Yellow ] (confidence: 0.625)
[DT_Colombo] --> [Tri_Yellow ] (confidence: 0.625)
[DT_Colombo] --> [CR_No , Tri_Yellow ] (confidence: 0.625)
[DT_Kalutara] --> [Tri_Yellow ] (confidence: 0.637)
[DT_Kalutara] --> [CR_No , Tri_Yellow ] (confidence: 0.637)
[S_Female] --> [CR_No , Tri_Yellow ] (confidence: 0.640)
[S_Female] --> [Tri_Yellow ] (confidence: 0.640)
[S_Female, DT_Gampaha] --> [CR_No , Tri_Yellow ] (confidence: 0.645)
[S_Female, DT_Gampaha] --> [Tri_Yellow ] (confidence: 0.646)
[DT_Colombo] --> [CR_No ] (confidence: 0.729)
[S_Male, DT_Gampaha] --> [CR_No ] (confidence: 0.734)
[S_Male] --> [CR_No ] (confidence: 0.740)
[DT_Gampaha] --> [CR_No ] (confidence: 0.757)
[S_Female, DT_Kalutara] --> [CR_No ] (confidence: 0.762)
[S_Female] --> [CR_No ] (confidence: 0.763)
[DT_Kalutara] --> [CR_No ] (confidence: 0.766)
[S_Female, DT_Gampaha] --> [CR_No ] (confidence: 0.776)
[CR_No , S_Male, DT_Gampaha] --> [Tri_Yellow ] (confidence: 0.819)
[CR_No , DT_Gampaha] --> [Tri_Yellow ] (confidence: 0.825)
[CR_No , S_Female, DT_Gampaha] --> [Tri_Yellow ] (confidence: 0.831)
[CR_No , S_Male] --> [Tri_Yellow ] (confidence: 0.831)
[CR_No , DT_Kalutara] --> [Tri_Yellow ] (confidence: 0.832)
[CR_No ] --> [Tri_Yellow ] (confidence: 0.835)
[CR_No , S_Female] --> [Tri_Yellow ] (confidence: 0.839)
[CR_No , DT_Colombo] --> [Tri_Yellow ] (confidence: 0.857)
[S_Female, CR_Yes ] --> [Tri_Orange ] (confidence: 0.911)
[CR_Yes ] --> [Tri_Orange ] (confidence: 0.916)
[S_Male, CR_Yes ] --> [Tri_Orange ] (confidence: 0.922)
[Tri_Yellow , S_Female, DT_Gampaha] --> [CR_No ] (confidence: 0.999)
[Tri_Yellow , DT_Gampaha] --> [CR_No ] (confidence: 0.999)
[Tri_Yellow , S_Female] --> [CR_No ] (confidence: 0.999)
[Tri_Yellow ] --> [CR_No ] (confidence: 1.000)
[Tri_Green ] --> [CR_No ] (confidence: 1.000)
[Tri_Orange ] --> [CR_Yes ] (confidence: 1.000)
[Tri_Yellow , S_Male] --> [CR_No ] (confidence: 1.000)
[Tri_Yellow , DT_Kalutara] --> [CR_No ] (confidence: 1.000)
[Tri_Yellow , DT_Colombo] --> [CR_No ] (confidence: 1.000)
```

```
[S_Female, Tri_Orange ] --> [CR_Yes ] (confidence: 1.000)
[S_Male, Tri_Orange ] --> [CR_Yes ] (confidence: 1.000)
[Tri_Yellow , S_Male, DT_Gampaha] --> [CR_No ] (confidence: 1.000)
```

6.2.13 Patient Group: Pediatrics (Age=<16)



168 patient cases utilize the ambulance services in Colombo District which out of 692 in pediatric cases.

Index	Nominal value	Absolute count	Fraction
1	Sepsis	76	0.452
2	ACR	37	0.220
3	Shock	25	0.149
4	Burn Injuries	19	0.113
5	Speial Patient Groups	11	0.065

Based on the above statistics the Colombo pediatrics have identified Sepsis in 76 cases out of 168. The health decision-makers should consider attending for these cases to handle in a proper way in the future.

Pediatrics: Gampaha District

321 patient cases utilize the ambulance services in Gampaha District which out of 692 in pediatric cases.

Index	Nominal value	Absolute count	Fraction
1	Sepsis	180	0.561
2	ACR	58	0.181
3	Burn Injuries	34	0.106
4	Shock	29	0.090
5	Speial Patient Groups	20	0.062

There are 181 cases identified for Sepsis under age 16 group out of 321 cases which is 56.1 % of pediatrics cases identified in sepsis which need special attention by the health authorities' in-order to find out the reasons in Gampaha Districts. The triage category is as below for the above cases.

Index	Nominal value	Absolute count	Fraction
1	Yellow	198	0.617
2	Orange	73	0.227
3	Green	43	0.134
4	Red	7	0.022

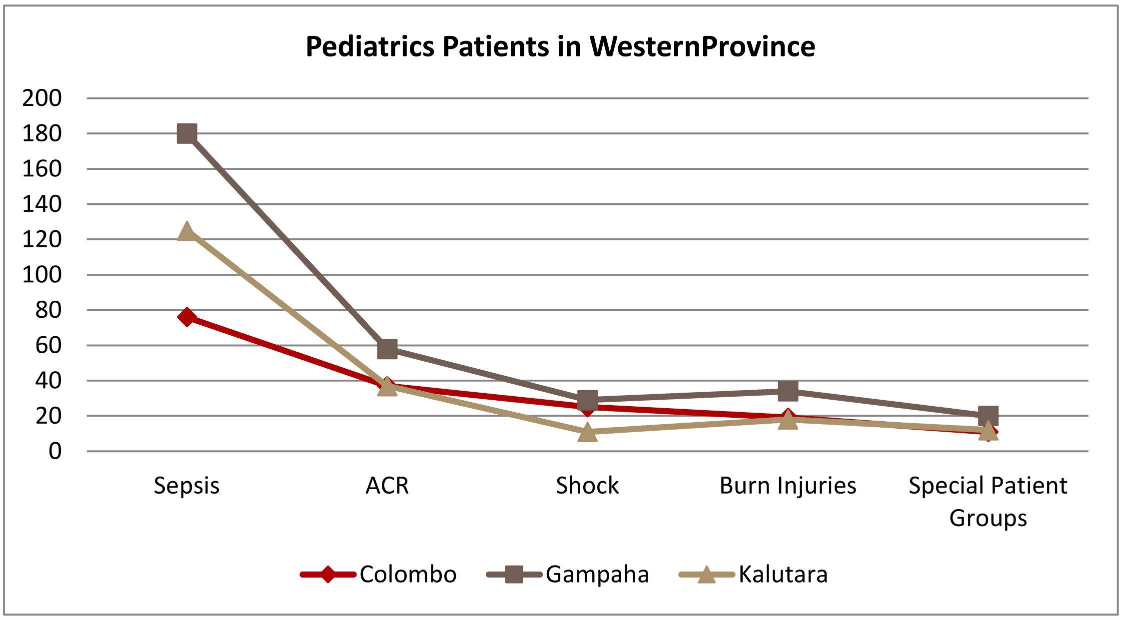
Pediatrics: Kalutara District

203 patient cases utilize the ambulance services in Kalutara District which out of 692 in pediatric cases.

Index	Nominal value	Absolute count	Fraction
1	Sepsis	125	0.616
2	ACR	37	0.182
3	Burn Injuries	18	0.089
4	Speial Patient Groups	12	0.059
5	Shock	11	0.054

Pediatrics in Western Province

Pediatrics	Sepsis	ACR	Shock	Burn Injuries	Special Patient Groups
Colombo	76	37	25	19	11
Gampaha	180	58	29	34	20
Kalutara	125	37	11	18	12



6.2.13 Patient Group: Age>16

Kalutara

Index	Nominal value	Absolute count	Fraction
1	Sepsis	663	0.542
2	ACR	213	0.174
3	Speial Patient Groups	124	0.101
4	Shock	114	0.093
5	Burn Injuries	109	0.089

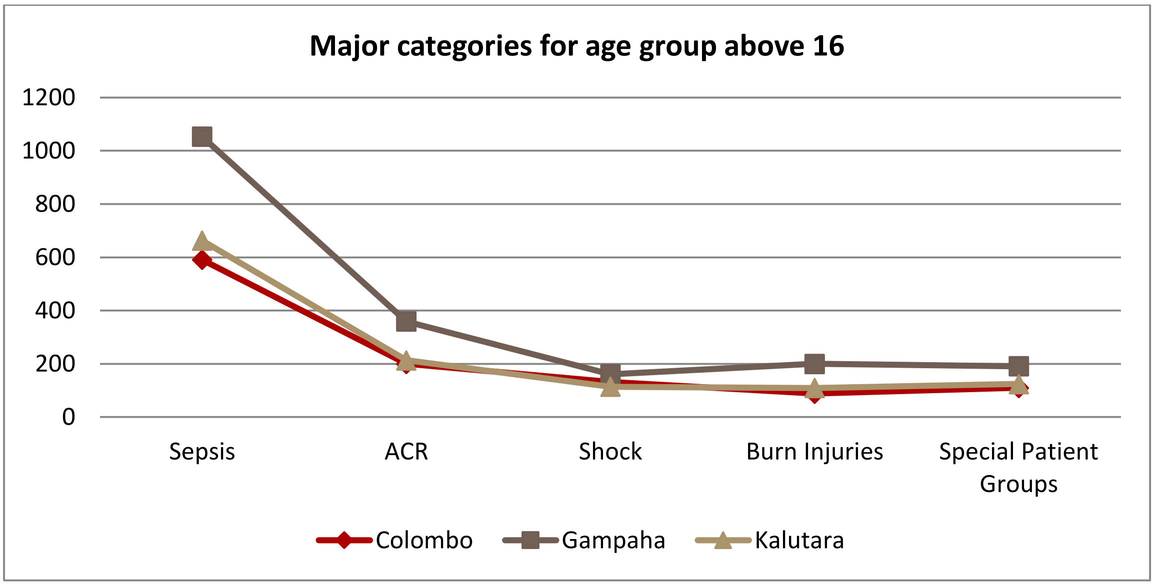
Gampaha

Index	Nominal value	Absolute count	Fraction
1	Sepsis	1053	0.536
2	ACR	359	0.183
3	Burn Injuries	200	0.102
4	Speial Patient Groups	191	0.097
5	Shock	161	0.082

Colombo

Index	Nominal value	Absolute count	Fraction
1	Sepsis	591	0.527
2	ACR	200	0.178
3	Shock	132	0.118
4	Speial Patient Groups	110	0.098
5	Burn Injuries	88	0.079

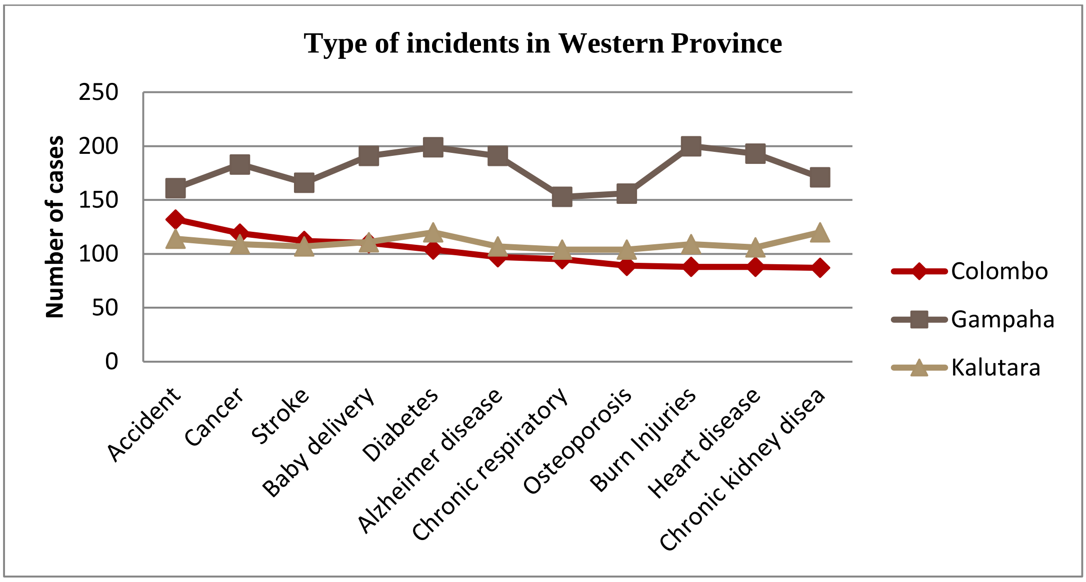
Age group above 16					
	Sepsis	ACR	Shock	Burn Injuries	Special Patient Groups
Colombo	591	200	132	88	110
Gampaha	1053	359	161	200	191
Kalutara	663	213	114	109	124



6.2.13 Type of Incidents in Western Province

Type of Incidents in Western Province

	Colombo	Gampaha	Kalutara
Accident	132	161	114
Cancer	119	183	109
Stroke	112	166	107
Baby delivery	110	191	111
Diabetes	104	199	120
Alzheimer disease	97	191	107
Chronic respiratory	95	153	104
Osteoporosis	89	156	104
Burn Injuries	88	200	109
Heart disease	88	193	106
Chronic kidney diseases	87	171	120

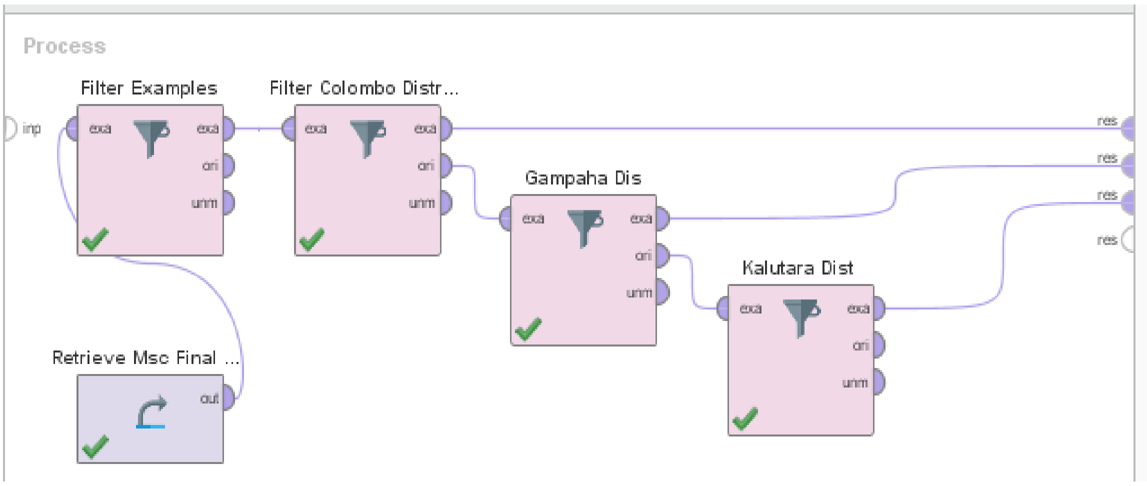


Triage category Status for the Western Province

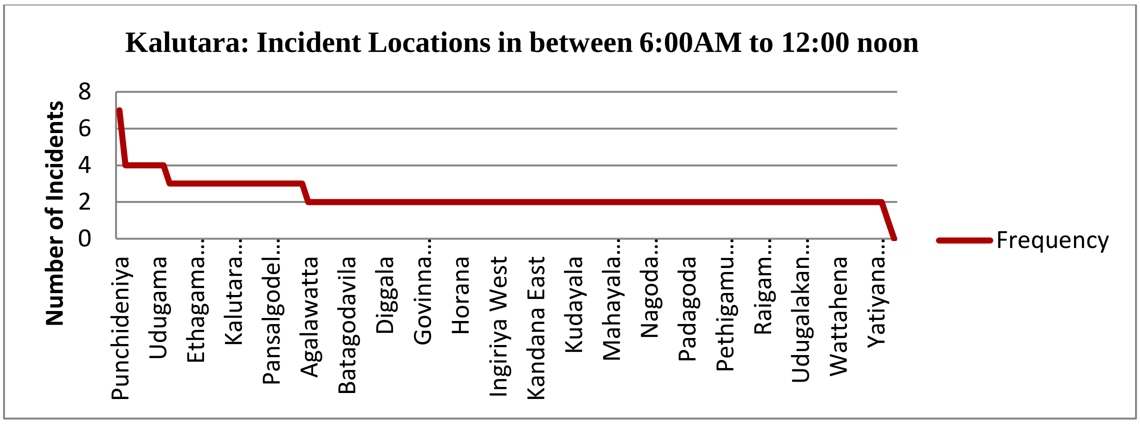
	Kalutara	Gampaha	Colombo
Yellow	769	1211	700
Orange	264	430	272
Green	158	256	112
Red	32	67	37

Type of cases between 06:00 AM to 12:00 Noon for the Western Province

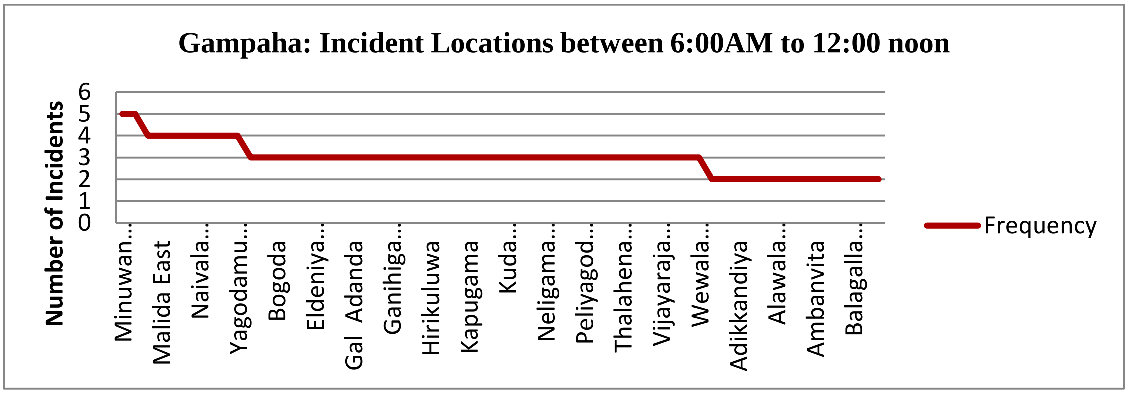
The process has been created to filter the data for the time interval of 6:00 AM to 12:00 Noon for the Districts of Colombo, Gampaha, and Kalutara separately as in below.



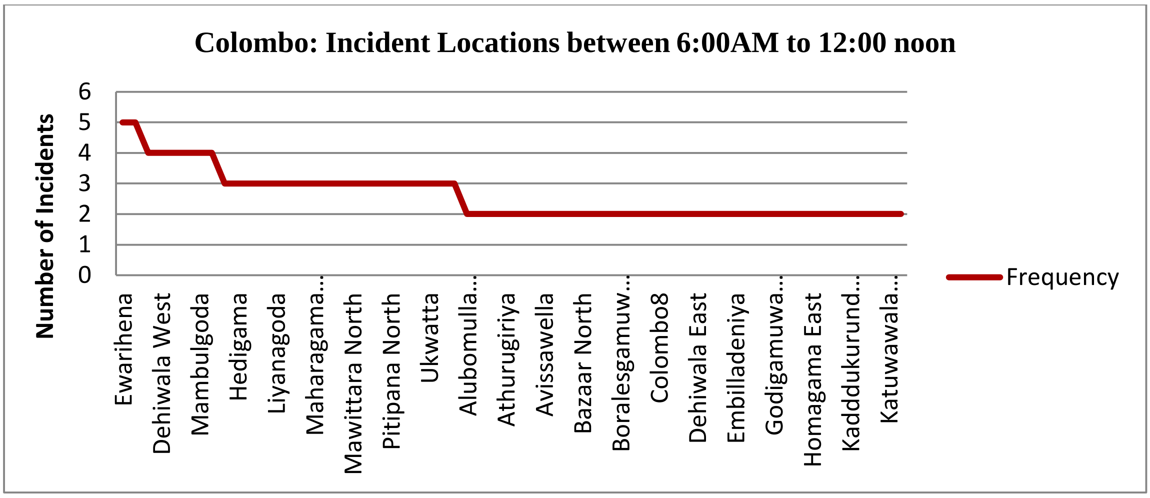
Based on the output of the process drafted the below graph for the clear vision of understanding.



Gampaha

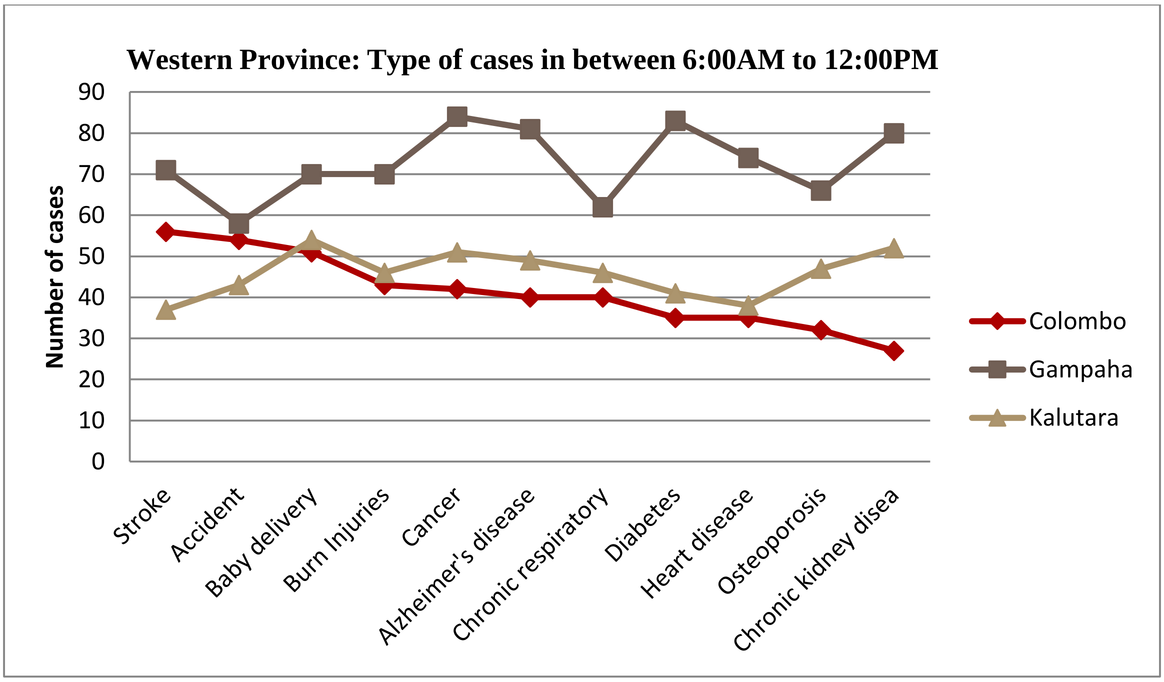


Colombo



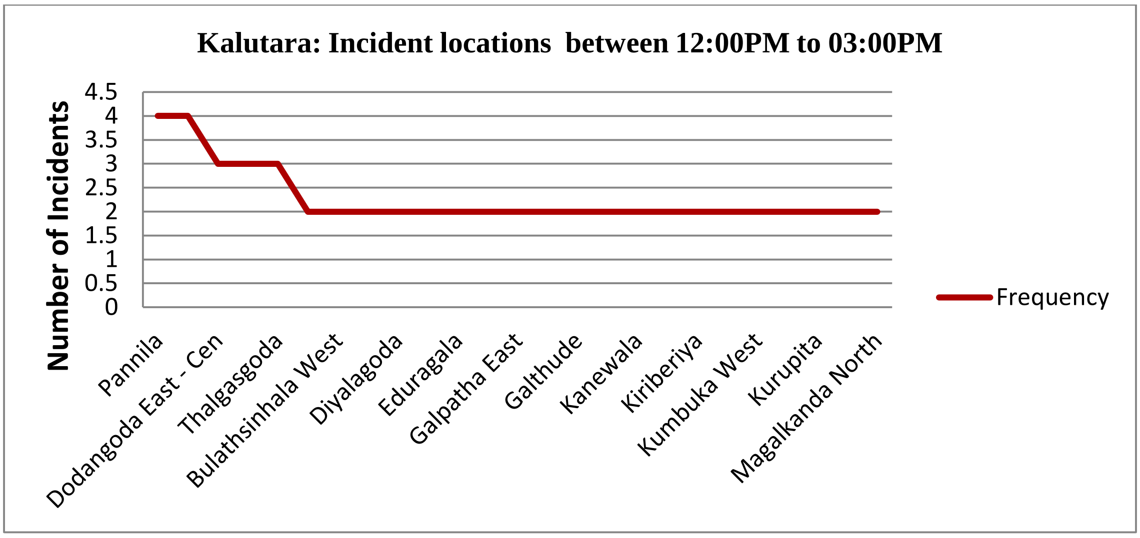
Type of cases between 6:00 AM to 12:00 PM for Districts.

	Colombo	Gampaha	Kalutara
Stroke	56	71	37
Accident	54	58	43
Baby delivery	51	70	54
Burn Injuries	43	70	46
Cancer	42	84	51
Alzheimer's disease	40	81	49
Chronic respiratory	40	62	46
Diabetes	35	83	41
Heart disease	35	74	38
Osteoporosis	32	66	47
Chronic kidney disease	27	80	52

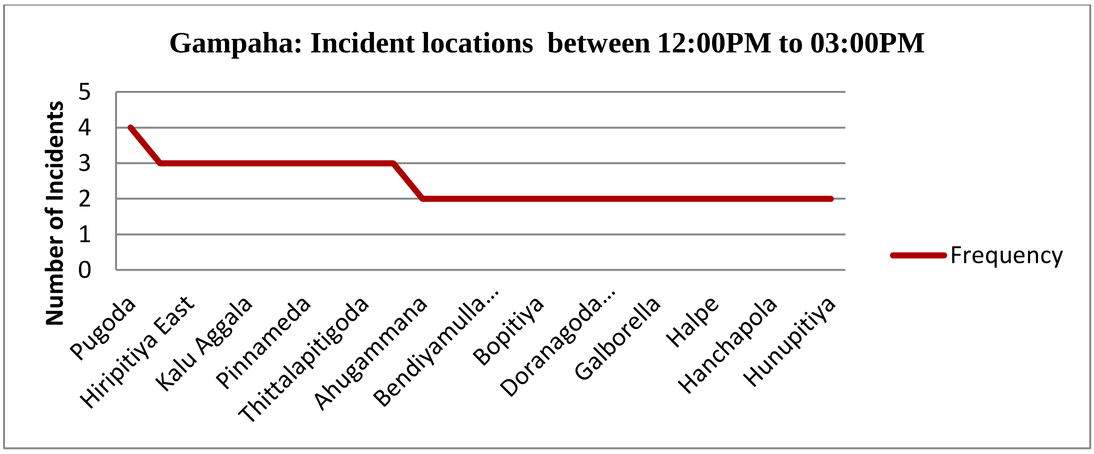


Type of cases between 12:00 Noon to 03:00 PM for the Western Province

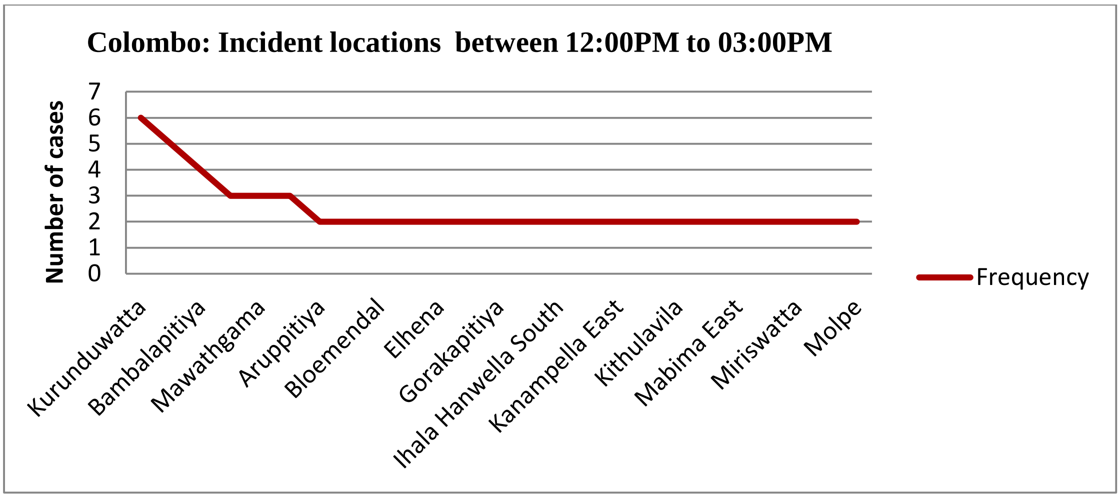
Kalutara: 244 cases



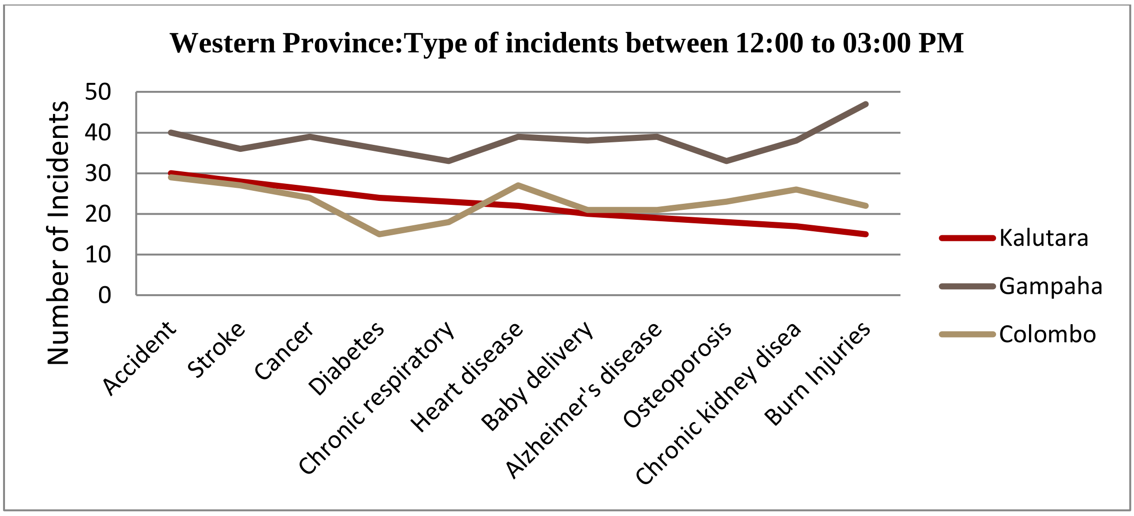
Gampaha: 419 cases



Colombo: 252 cases

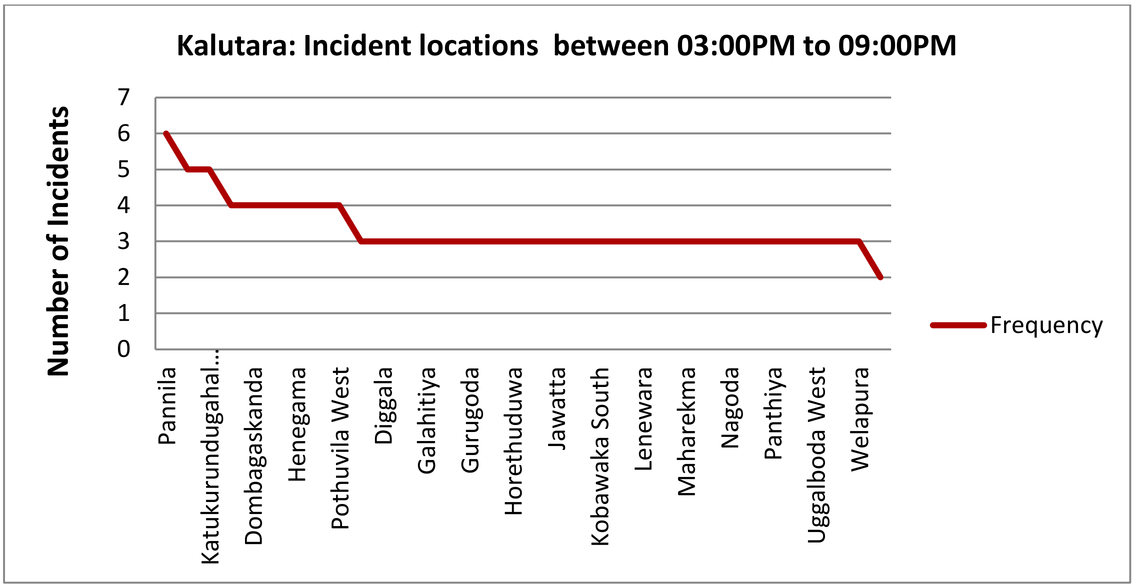


Type of Incidents between 12:00 to 3:00 PM

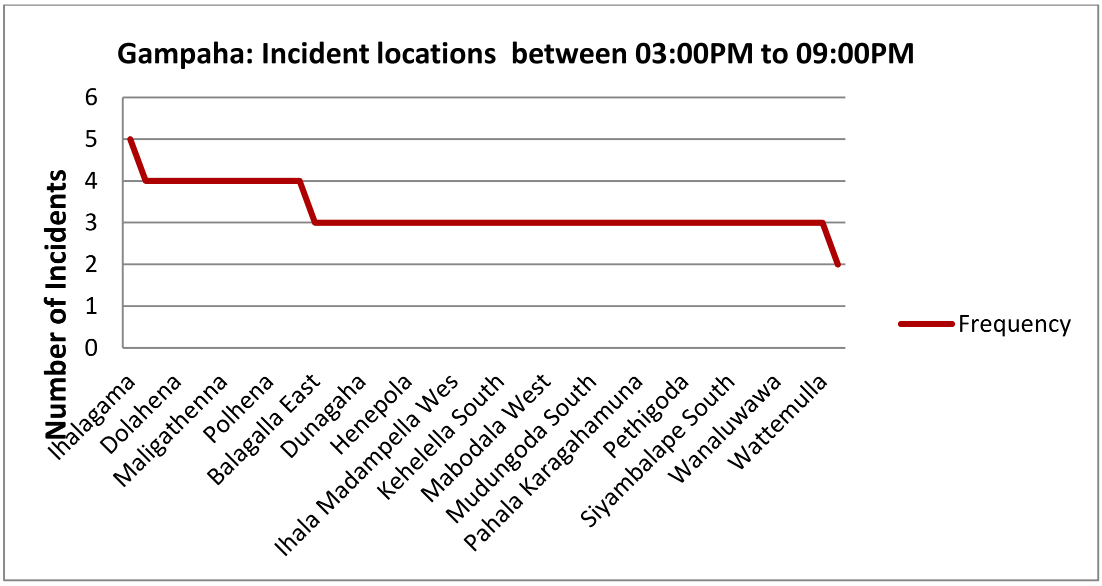


Type of cases between 03:00 PM to 09:00 PM for the Western Province

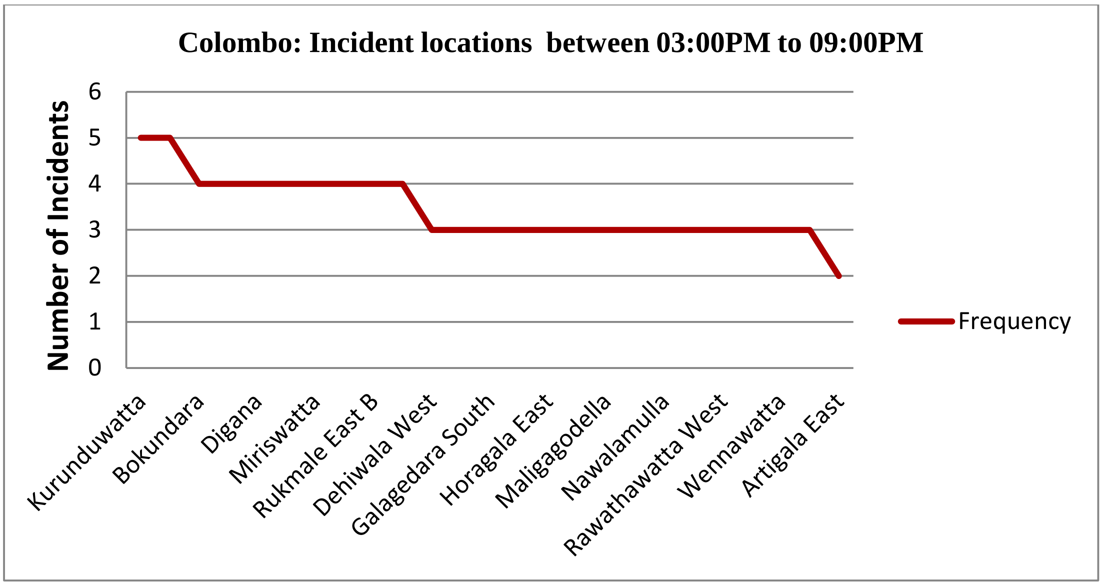
Kalutara: 483 cases



Gampaha: 728 Cases

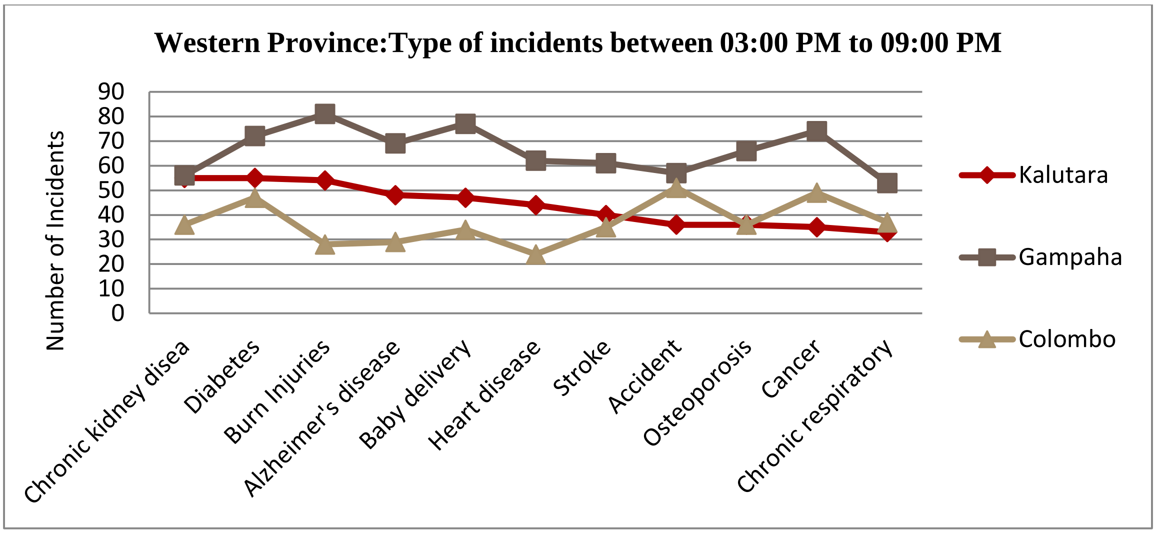


Colombo: 406 Cases



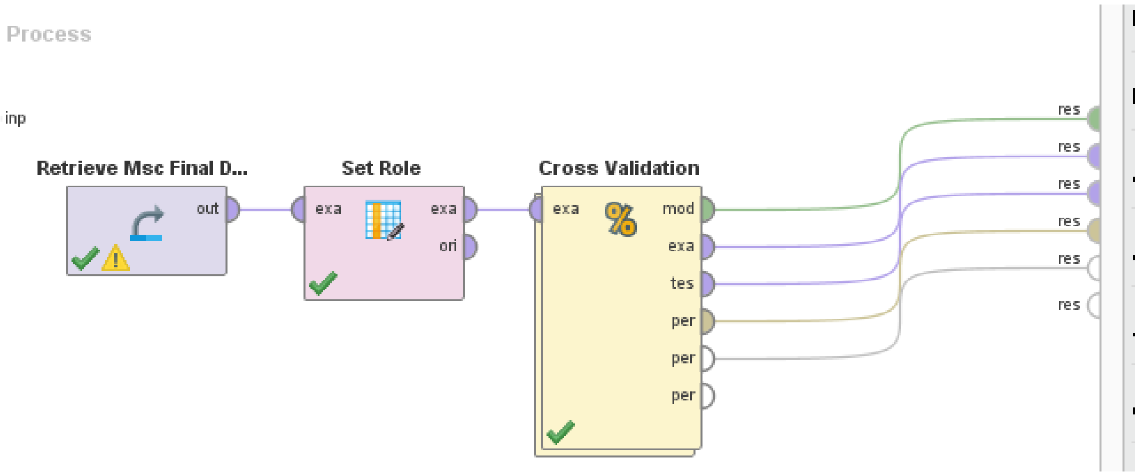
Western Province: Type of incidents Summary

Incident type	Kalutara	Gampaha	Colombo
Chronic kidney disease	55	56	36
Diabetes	55	72	47
Burn Injuries	54	81	28
Alzheimer's disease	48	69	29
Baby delivery	47	77	34
Heart disease	44	62	24
Stroke	40	61	35
Accident	36	57	51
Osteoporosis	36	66	36
Cancer	35	74	49
Chronic respiratory	33	53	37

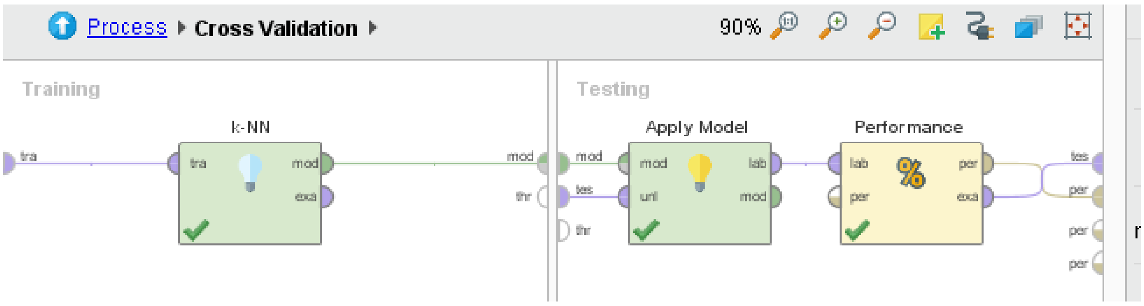

Western Province:Type of incidents between 03:00 PM to 09:00 PM
Number of Incidents
90
80
70
60
50
40
30
20
10
0
Chronic kidney disea
Diabetes
Burn Injuries
Alzheimer's disease
Baby delivery
Heart disease
Stroke
Accident
Osteoporosis
Cancer
Chronic respiratory
Kalutara
Gampaha
Colombo

Appendixture-2

KNN Classification model



The cross-validation setup for the above process



Performance vector of the KNN Classification model

ExampleSet (Set Role)   KNNClassification (k-NN)

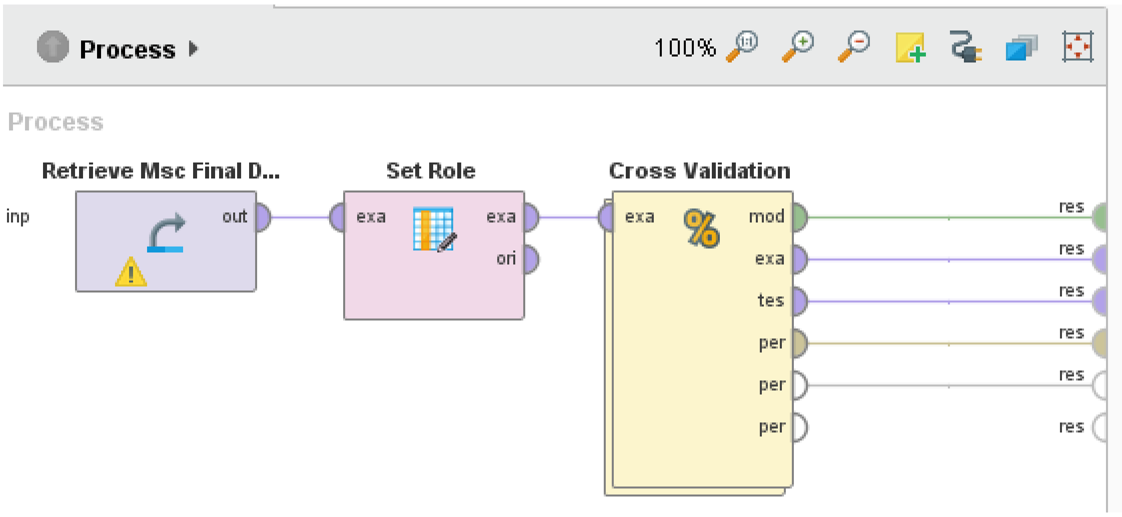
PerformanceVector (Performance)

Table View   Plot View

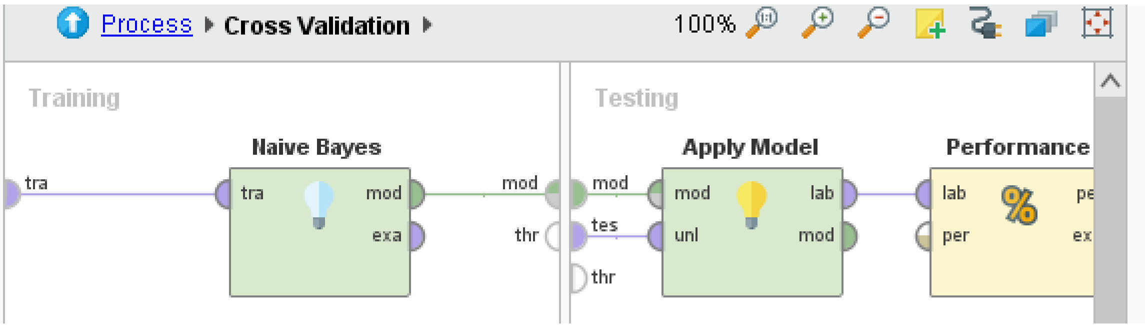
accuracy: 91.44% +/- 0.47% (micro average: 91.44%)

	true Yellow	true Orange	true Red	true Green	class precision
pred. Yellow	3089	57	5	207	91.99%
pred. Orange	4	1071	98	0	91.30%
pred. Red	0	1	0	0	0.00%
pred. Green	50	6	0	412	88.03%
class recall	98.28%	94.36%	0.00%	66.56%	

Naïve Bayes Classification



The cross-validation setup for the above process.



The output performance for the Naïve Bayes Classification

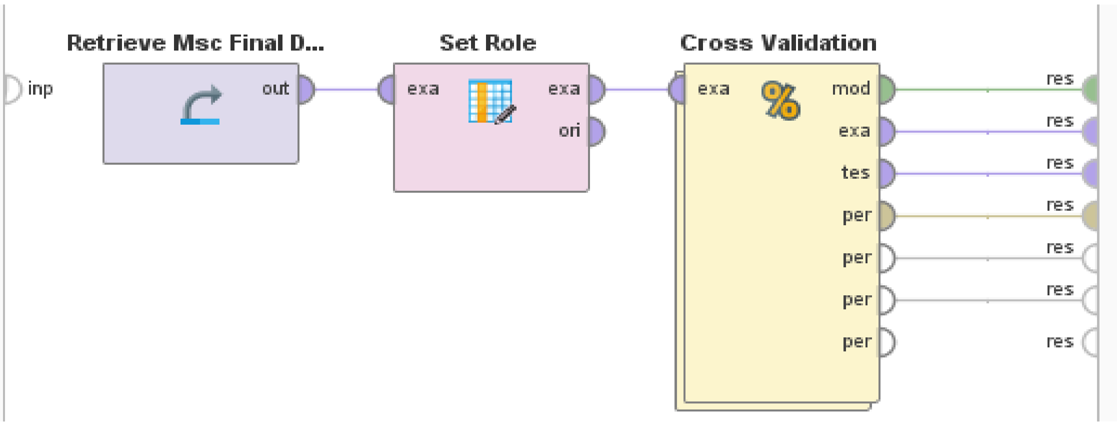
ExampleSet (Set Role) | SimpleDistribution (Naive Bayes) | PerformanceVector (Performance)

Table View | Plot View

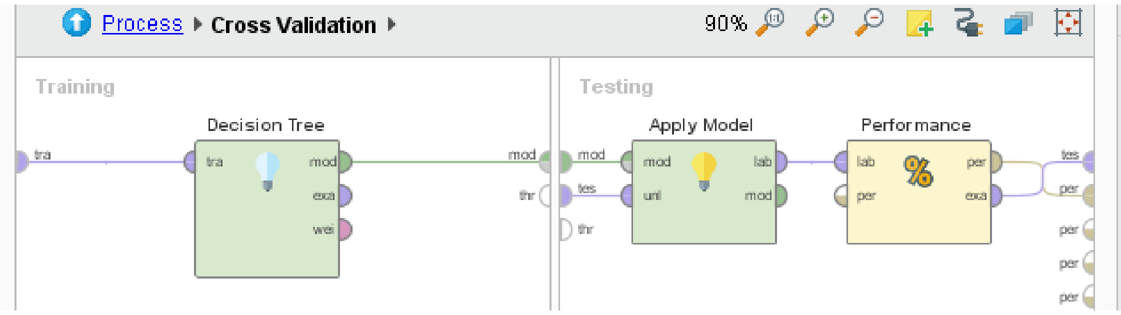
accuracy: 84.14% +/- 1.30% (micro average: 84.14%)

	true Yellow	true Orange	true Red	true Green	class precision
pred. Yellow	2949	1	0	389	88.32%
pred. Orange	0	1009	84	0	92.31%
pred. Red	0	125	19	0	13.19%
pred. Green	194	0	0	230	54.25%
class recall	93.83%	88.90%	18.45%	37.16%	

6.3.3 Decision Tree Classification



The cross-validation model for the above setup



The output of the performance.

ion) | ExampleSet (Set Role) | Tree (Decision Tree)

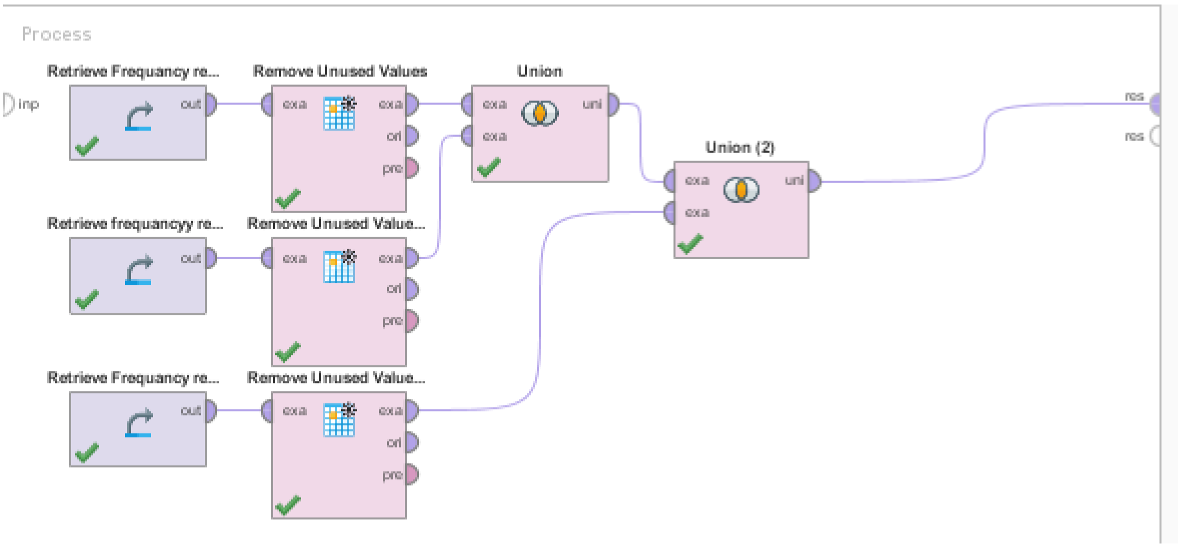
PerformanceVector (Performance)

Table View Plot View

accuracy: 92.64% +/- 0.56% (micro average: 92.64%)

	true Yellow	true Orange	true Red	true Green	class precision
pred. Yellow	3092	0	0	207	93.73%
pred. Orange	1	1126	101	0	91.69%
pred. Red	0	9	2	0	18.18%
pred. Green	50	0	0	412	89.18%
class recall	98.38%	99.21%	1.94%	66.56%	

Incident locations in frequency based



The output of the above process.

Row No.	District	Incient_Loc...	Incident_type	Triage Cate...	Row No	Frequency ↓
257	Gampaha	Bothale Paha...	Cancer	Orange	59	6
309	Gampaha	Gonahena E...	Heart disease	Orange	19	6
437	Gampaha	Opathella	Stroke	Orange	141	6
596	Kalutara	Govinna North	Heart disease	Orange	158	6
17	Colombo	Bambalapitiya	Accident	Orange	35	5
29	Colombo	Bopetta	Heart disease	Orange	311	5
201	Colombo	Udagama	Stroke	Orange	72	5
262	Gampaha	Dalupitiya East	Stroke	Orange	131	5
285	Gampaha	Eldeniya East	Stroke	Orange	171	5
349	Gampaha	Kadawala No...	Heart disease	Orange	295	5
433	Gampaha	Neligama	Stroke	Orange	351	5
475	Gampaha	Polwatta	Heart disease	Orange	204	5
500	Gampaha	Thalahena	Accident	Orange	47	5
16	Colombo	Avissawella	Stroke	Orange	104	4
30	Colombo	Boralesgamu	Osteoporosis	Orange	51	4