2021

A Fuzzy expert based multi-criteria performance appraisal model for police officers.

Egessa, Bryan
School of Computing and Engineering Sciences
Strathmore University

A Fuzzy Expert Based Multi-Criteria Performance Appraisal Model for Police Officers

BRYAN EGESSA
090524

Submitted to the School of Computing and Engineering Sciences in partial fulfilment of the requirements for the Degree of Master of Science in Computer Based Information Systems at Strathmore University

School of Computing and Engineering Sciences
Strathmore University

April 2021

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Student Name: **Bryan Were Egessa**

Signature: ……

Date: 20th September 2021

Approval

The dissertation of Bryan Were Egessa was reviewed and approved for examination by the following:

Dr. Joseph Orero, PhD
School of Computing and Engineering Sciences
Strathmore University

Dr. Bernard Shibwabo
Director of Graduate Studies
Strathmore University

Abstract

Performance appraisals play a key role in management of human capital and in particular in determining the effectiveness of an employee and to map out their developmental needs. Subjective appraisals are prevalent especially in the public sector, which has contributed to bias and inconsistent ratings and lack of transparency in determining promotions and reward. The appraised employees therefore hardly expect benefits from the exercise other than organisational compliance with regulations.

Effective performance appraisal should be void of bias with measurable outputs that can guide various human capital management procedures. The purpose of this research was to establish the gap in appraisal practices in policing organisations and to present a solution that would enhance objectivity in the appraisal of police officers. Recommendations from previous research point to the need to have appraisal ratings arrived at from multiple sources that would include the public who are the primary consumers of policing services. Police satisfaction surveys and needs analysis reports have yielded recommendations that the public would provide valuable input in determining the effectiveness of a police officer. Other key players in determining appraisal ratings are the officer's supervisor and peers with whom they have handled assignments.

This research applied the prototyping methodology to design, develop and test a fuzzy expert based appraisal model with multiple appraisal data inputs and a single performance score output. To determine performance scores through a uniform procedure, a fuzzy controller was used to approximate the relationship between inputs and outputs via interpolation. The reason for applying fuzzy logic for the development of the appraisal model lies in the fact that vagueness is expected whenever human decisions are made. The crisp data inputs to the fuzzy controller were obtained from a mobile application developed to capture incident reports in real-time from the public, enable swift response by an officer and enable user rating upon completion of the assignment. Fuzzy logic is effective where multi-criteria decision-making is required such as an appraisal of employee performance. The web application provides an input interface for ratings by supervisors, peers and sub-ordinates as well as administrative tools. This model will enhance objectivity in the appraisal of police officers, which will also enable accurate identification of an individual's development needs.

Dedication

I dedicate this dissertation to my parents Philip and Phenike, my siblings Laura, Kevin and Regina who encouraged and supported me through this journey. My godsons Adrian and Damian, you were a source of motivation. May the Almighty God be glorified for His grace and provision.

Acknowledgements

I am grateful to God for granting the opportunity, health and resources that have enabled me to pursue this course.

I would like to appreciate my supervisor Dr. Joseph Orero. Through his guidance and support, I have been inspired to focus on delivering quality. My appreciation also goes to Dr. Vincent Omwenga for his valuable contribution to this study.

My parents, siblings and godsons, you have been supportive through your moral and material support. The strength I have drawn from your prayers is immeasurable and for this, I am truly thankful to each one of you.

Finally, I extend my thanks to the entire Strathmore University community for creating an enabling environment and general support towards achieving this Master's degree.

Table of Contents

List of Figures

List of Tables

List of Equations

Abbreviations

AI	-	Artificial Intelligence
CAW	-	Cumulative Average Weighting
FAM	-	Fuzzy Associative Matrix
FIS	-	Fuzzy Inference System
KDD	-	Knowledge Discovery in Database
MCDM	-	Multi Criteria Decision Making
MBO	-	Management by Objectives
ML	-	Machine Learning
NPS	-	National Police Service
RSPS	-	Royal Swaziland Police Service
SVM	-	Support Vector Machine
UML	-	Unified Modelling Language

Definition of Terms

Gazetted Officer - a police officer holding the rank of a superintendent and above (Ransley, 2009).

Performance appraisal (PA) - the measurement of work and its results by using the scale and index that we can measure the desired quantity and quality with precision and free of personal judgments and vague criteria of evaluation (Ali et al., 2012).

Performance management (PM) - is a process whereby goals are directed toward ensuring that organizational processes are in place to maximize the productivity of employees and the organization. It involves measuring and improving the value of the workforce (Keenan, 2015).

Chapter 1: Introduction

1.1 Background of Study

Law enforcement agencies perform a wide range of functions that include maintaining social order and building effective partnerships with communities. Knowing how these functions are performed is important as it has a direct effect on evaluating the style of policing (Jahan, 2014). Consequently, the performance of individual officers with regard to performance of these functions needs to be monitored. This ensures that the officers maintain the highest standards of ethics and professionalism in the discharge of their duties and tracks their professional development. The introduction of performance management is consistent with a desire to improve public trust by improving service provision, accountability and justifies budget requests (Bianchi, 2010). According to Politt (2013), professionalism in policing should be enhanced to bring police officers at par with professionals in other fields. This can be done by ensuring that the performance of Police Officers is evaluated impartially and through a standard procedure.

In the developed democracies (Bennet, 2019), policing agencies are viewed as a form of enterprise that are funded by the public who in turn should receive a return on their investment through quality service delivery. Various jurisdictions have formulated methodologies that can be utilised in making comprehensive assessments of performance. Experts in the field of policing agree that performance measurement systems should capture the expectations from an enlightened society with a view to holding individual officers accountable for their actions. In the United States, a balanced scorecard approach has been used to ensure that measurement schemes capture the diverse policing outcomes and ensure that the quality of engagement with the public is taken into account (Bennet, 2019).

Most African countries inherited a policing culture that was moulded around crushing dissent during the colonial period. However, there has been a gradual shift towards international best practice of policing as a service, which has as an objective the measurement of police performance to ensure improvement of effectiveness and officer conduct (Batts et al., 2012). These measurement systems heavily rely upon accurate incident recording. The resultant systems have been found to work best where there is buy-in and where metrics are developed with a focus of what is in control of the police. The demand from the public is that the police will enforce law, reduce incidents of crime and treat the

people they encounter professionally and fairly. The performance indicators developed so far may be few but result in valid and reliable measures (Berger, 2019).

In Kenya, the Public Service Commission approved a performance appraisal instrument that applies to all public servants, the Police Service included (PSC 37A, 2016). Correct use of this instrument would enhance performance while motivating officers to achieve set targets that can then be linked to promotions to positions with heavier responsibilities and help in strategic deployment of performers. According to Ransley (2009), in determining the suitability of a Police officer for promotion, performance appraisals should be taken into account. The areas of appraisal include courses attended, achievement of performance targets, values and competencies, work ethic and discipline, qualities, leadership skills and disciplinary record. Measurement of these parameters however rests fully at the discretion of the appraising supervisor.

The National Task Force on Police Reforms noted that if properly applied, an appraisal instrument would be a major step towards ensuring individual Police Officers remain accountable as they discharge their duties. This study introduces the multi-criteria evaluation of police performance in Kenya where the public who are the consumers of policing services have an input in rating the services provided. The field officers too will get a platform through which they can demonstrate their competence in handling their customers. The development of the appraisal prototype is expected to enhance objectivity in the appraisal of police officers.

1.2 Problem Statement

Surveys reveal widespread dissatisfaction with the performance appraisal process (Jones, 2018). Despite government efforts in reforming the security sector, policing standards have continued to deteriorate; criminal activities have increased and are more sophisticated (Barasa, 2017). This has been largely attributed to lack of measurable outputs to be used in determining the promotion and deployment of Police Officers. These functions should take into account the individual officer's appraisal ratings. Deployment of a non-performing officer to a position of heavy responsibility waters down policing standards with serious implications on national security.

There have been concerns that the appraisal ratings from current appraisal systems are determined purely at the discretion of a single source - the immediate supervising senior ranking officer and may be open to bias and personal opinions. (Janssens, et al., 2017). An effective performance appraisal system focuses on performance variables as opposed to personal traits and the results must be acceptable by all concerned parties (Hassan et al., 2013). Police performance systems have problems such as data manipulation and subjectivity in ratings by supervisors the focus should shift from abstract management styles to data driven and precision-based methods (Pasha, 2014)

To address the problem of partiality in appraisal, an objective and accountable appraisal model that will apply the 360-degree method, taking into account feedback from the consumers of Police services was proposed to reduce bias that may result from ratings derived from a single source. The proposed model is expected to transform the Police institutional culture from a rules-bound to a results-driven system through continuous feedback.

1.3 Research Objectives

1.3.1 General Objective

The purpose of this study is to develop a multi-rater appraisal model that incorporates feedback from the consumers of policing services, their supervisors and their peers for enhanced objectivity in the appraisal process.

1.3.2 Specific Objectives

i. To review existing models used in appraisal in Police agencies.
ii. To examine the algorithms that can be used in performance appraisal systems
iii. To develop an appraisal prototype based on the assessed models and algorithms.
iv. To validate the prototype.

1.3.3 Research Questions

i. Which models have been used in appraisal by Police agencies?
ii. What algorithms can be used in performance appraisals?
iii. How will the proposed tool be designed and developed?
iv. How will the tool's functionalities be tested?

1.4 Justification

Performance measurement can improve citizen trust in government directly through citizen participation in the evaluation process or indirectly by improving citizens' perceptions of government performance (Yang and Holtzer, 2006). A needs analysis report of the Kenya Police Service recommended establishment of a system that would effectively monitor performance of police officers. Clients served may offer a different perspective on individual performance in service delivery. This system could include citizen feedback as part of their performance appraisal (Transparency International Kenya, 2016)

The proposed system should influence the way in which police officers carry out their duties, which will have a positive impact on security as well as raise the level of quality of service rendered since it will reduce subjectivity associated with the traditional appraisal systems. It is highly likely that employees who are dissatisfied with appraisals will negatively affect the organization's operations (Walker, 2015). The system will allow officers an opportunity to have their efforts recognised through the feedback of the people they serve. An information system would also ensure that every reported incident is documented and acted upon. There have been concerns that some offences are never booked as they are either too difficult to deal with or would paint regional commanders in a bad light as far as their crime prevention strategies are concerned. Fuzzy logic mimics human behaviour and is ideal in multi-criteria decision making such as in performance appraisal. The results of this research will also be beneficial in narrowing the gap in existing literature on modern techniques of appraising performance of police officers.

1.5 Scope and Limitation

This study aims at developing a fuzzy expert based information system that would refine the appraisal process of police officers against predetermined institutional goals. The research was limited to having reported incidents as a basis for customer evaluation. In order to ensure completion of the research within the predetermined timelines while maintaining its significance, Matlab was used to develop the Fuzzy Rule Base.

1.6 Limitations

The application front end was developed for use on the android platform. The study limits itself to the Cumulative Average Weighting algorithm to determine cumulative appraisal score from the members of the public (customers) and Fuzzy Logic for performance score computation. This research was also limited by time constraints as per the academic research requirements.

1.7 Assumptions

An assumption was made on the self-appraisal component of the 360-degree appraisal technique. It was assumed that the officer's report back to the system, through the mobile application, following attendance to an incident would be a self-assessment. Another assumption was made on the validation on the proposed model. It was assumed that the respondents engaged in the validation process understood the questions and that they would provide unbiased responses regarding the practicality of the model.

Chapter 2: Literature Review

2.1 Introduction

In this chapter, performance appraisal systems and procedures were reviewed to gain insight into current practice. Further, systems and models applicable specifically to Police institutions across various jurisdictions were reviewed to identify gaps and challenges in rating the performance of law enforcers. Algorithms and models suitable for use in performance rating were investigated to determine suitability for deployment in the proposed improved solution's conceptual framework.

2.2 The Concept of Performance Appraisal in Policing Agencies

The effectiveness of a law enforcement body's performance appraisal system is a requirement that will guarantee the success of its strategic deployment practices. Rapid and effective re-alignment ensures restoration or maintenance of order in the command structure. Performance appraisal is generally one of the most complex components of human resource management (Gul & O'Connel, 2013). The process creates a destructive influence on the involved parties because current organisational practice creates room for dissatisfaction with the methods used (Bernardin & Wiatrowski, 2013).

There are varied views on best practice; whether performance should be measured in terms of the results produced by employees or in terms of work-related behaviours. However, there is consensus that measuring personal traits has several drawbacks (Migiro & Taderera, 2011). According to Wan et al (2012), performance appraisal processes can be discredited due to the bias that is often associated with the performance reviews. Indeed, as Weick (2001) asserts, it is common to have reviews that are discredited by other reviewers. Most review systems thus require a redundant mechanism of checks and measures that ensure checking and re-checking of the performance reports as they are presented. It also demands that the process be developed in a transparent manner that does not discriminate on employees based on their natural weaknesses or differences. Migiro and Taderera (2011) argues that it is important to get the actual performance report of the individual as it presents a true measure of the person's ability to perform tasks within an organization.

Appraisal of employees in an organization requires significant investment in data collection and research methods. The modern approach of performance appraisal is centred on the employee as an individual. Performance appraisal serves not only as a motivator but also in employee development

and accurate formulation of staffing plans (Nair & Salleh, 2015). Employees with promising abilities and expertise can be assigned greater responsibility while those with lower can be recommended for training and mentorship. The employee and employer can as a result collaborate to create a win-win situation (Schultz, 2015). Performance evaluations and incentives in law enforcement should take into consideration performance measures significantly more complex so as not to create room for frustration on the part of officers and lead to a culture where accountability for performance is not taken into account (O'Toole Jr & Meier, 2014).

For an appraisal system to be effective, employees must believe that they have an opportunity for meaningful input into the appraisal process (Weick, 2001). As noted by Gichuhi et al. (2013), without the perception of fairness, a system that is designed to appraise, reward, motivate and develop can actually have the opposite effect and create frustration and resentment. Therefore, for the Police Service, the individual officer's input into their appraisal should be deduced from the quality of work they put in to fulfil the organisation's mandate.

2.3 Performance Appraisal Techniques

Employees play an important role in ensuring an organisation's strategic objectives are achieved. The linkage between the appraisal system used and the organization's management strategy ensures achievement of the organization's vision. Appraisals therefore have a role in communicating values, promoting flexibility and maximizing individual potentials and contributions. Performance appraisal activities are therefore important elements in human capital management (Walker, 2015).

Several performance measurement methods have been put in place by organisations and in some a combination of systems are used in arriving at a measure of an employee's performance. The choice method depends on the strategic objectives and the nature of work performed by individual employees. This study focussed on Multi-Source Feedback, Management by Objectives, Critical Incidents Method and Graphic Rating Scales.

2.3.1 Multi-Source Feedback

This method involves systematic collection of performance data on an employee from several sources of individuals such as supervisors, peers and subordinates and from different contexts such as customers. This data provides a broader perspective on an employee's competencies as well as

increases their self-awareness through feedback, which is associated with effective leadership and higher performance outputs (Good & Coombe, 2009).

Feedback from multiple sources leads to greater reliability on the appraisal scores, higher acceptance from the employees and provides specific behaviour based output. This enhances individual performance through identification of individual and professional strengths and weaknesses when done properly. Multisource feedback can be a valid process to support professional development, performance enhancement and leadership development.

The main drawback in this method is the difficulty in giving and receiving feedback as well as various rating errors. Mental distress may occur when feedback conflicts with an existing belief regarding personal performance within the evaluated individual. It is also criticised as being over reliant on generalised individual traits and of being based on incomplete description of past performance. However, these conflicts could be reframed as opportunities for improvement, determining training needs and capitalising on previously unknown strengths. Coaching of respondents on expectations from the evaluation has shown to greatly improve on the accuracy in reporting (Graham & Beuthin, 2018).

Multisource feedback is therefore a developmental tool in as much as it is evaluative since it leads to self-development, helps in aligning individual behaviour as well as performance with organisational values. It also enables introduction of organisational culture change, encourages positive attitude towards learning, fosters participative leadership and a focus on customer service, which all lead to organisational effectiveness by calling to attention important performance dimensions overlooked by the organisation. Organisations that have embraced multi-source feedback have reported improved employee relations and fair reward decisions (Ermongkonchai, 2008).

2.3.2 Management by Objectives

Management by objectives is a method used to define and measure performance of employees where importance is placed upon contribution to the organisation's goals. It is a process that converts organisational objectives to individual objectives. These objectives must be clearly communicated which results whose attainment leads to satisfaction of employees. Performance is generally defined in terms of measurable outcomes by comparing the delivered performance with expectations defined as objectives. This method bears several components, the first of which involves both the supervisor

and subordinate setting goals by defining goals and performance measurements for the subordinate. They determine the deliverables and how they will be measured along set timelines. Then follows objective feedback on milestones achieved in the process of working towards set goals (Cintron & Flaniken, 2011).

The system has the advantage of enabling better planning and improved motivation since expected results are known beforehand, fair evaluations since evaluations are made on the basis of results rather than personality, improved commitment through participatory goal setting, improved supervisory skills in areas of communication and counselling (Black et al., 2019).

Success of the system relies upon good communication between superiors and subordinates and defines the areas of responsibility for each party. Management by objectives is closely related to the goal-setting theory of motivation. Motivation and satisfaction of employees has been found to be higher when specific goals are set, when the goals are challenging but accepted and when there is feedback on performance. The system is characterised by a reward and sanction system that seeks to motivate individual productivity. MBO is easy to implement and keeps employees engaged since they have a clear understanding of their roles and responsibilities, which encourages self-direction rather than micro management (Islami et al., 2018).

However, the system has received criticism for it creates a lot of paperwork since emphasis is on quantitative rather than qualitative goals. At times, pressure to achieve results in autocratic implementation that goes against the motivation theory that advocates for acceptance for effectiveness (Black et al., 2019).

2.3.3 Critical Incidents Method

This method requires keeping written records on certain critical behaviours of employees that make the difference in their performance which are noted by their superiors as and when they occur. A critical incident is an important employee action that can either improve or hurt performance. The event has to have an impact on the department or organisation to merit consideration as a critical event. The evaluator rates employees based on critical events and how the employee reacted in such events. It aims to monitor employee performance and behavioural patterns during those incidents. The supervisor makes a note of both the positive and the negative aspects of the employee and objectively comments upon the employee's performance during such critical incidents (Saxena, 2010).

This approach places heavy emphasis on the study of a worker's behaviour and helps establish key requirements, which have been responsible for outstandingly effective or definitely unsatisfactory performance of a job or activity through direct observation. This observation may be crucial in separating assumptions or perception to reality as it brings to the fore what exactly an employee does on the job (Saxena, 2010). The use of critical incidents allows managers to make an objective review over the entire appraisal period and to coach employees on a continuous basis to ensure development. With clear standards and coaching, managers can minimise disagreements over ratings because the employees already know how they have performed based on their actions (Rao, 2004).

The major advantage is that ratings are based on actual observed events over the entire evaluation period and revisited at prearranged times. It minimises recency biases and encourages feedback towards future performance improvement (Chan, 2006). The method has drawn criticism since it is difficult to comprehensively report on every employee consistently. It is not effective since individual differences dictate that employees perform and react differently in the same circumstances. Revisiting past negative incidents after a time lapse may elicit negative emotions from employees (Saxena, 2010). A common error is the tendency to focus on the negative actions of employees since a good evaluation includes both positive and negative aspects. Some supervisors may focus on negative occurrences whose feedback may demoralise employees (Chan, 2006).

2.3.4 Graphic Rating Scales

The Graphic Rating Scale is a numerical scale that lists a number of traits and a range of performance for each after which an employee is rated by identifying the score that best describes their level of performance. The scale is developed along appropriate factors such as personality traits, job behaviour and outcomes relevant to the organisation's needs and lists a number of work dimensions and a performance range for each (Aggarwal & Thakur, 2013).

The trait scales are easy to develop and the method relatively inexpensive to implement and the results are easy to quantify. The rating scales are developed with the assumption that the evaluator is tuned towards objectivity and precision in reporting. Since the rating scales are standardised, it is easy to make comparisons across individuals. It is less time consuming and allows for quantitative analysis and comparison. The system appraises individuals based on previously established performance

dimensions therefore focusing on the past rather than the future. Supervisors conduct the appraisals with little input from employees (Majid, 2006).

This method is ideal where simplicity and quick evaluations are desired to make comparisons of employee performances. Some employers make provision for the evaluator to comment on the evaluation for each of the factors to enhance objectivity at arrival of scores and to provide justification for extreme scores. This aids in correcting shortcomings in judgement and ensures the rating process is handled professionally. This method requires consistent follow up of each employee by the supervisors, which may greatly alter the quality of eventual scores. An extreme score such as outstanding may only be awarded to employees who consistently perform beyond expectations and not just in a few instances. The system is easy to implement, however it cannot be used to provide guidance needed for improving performance (Cintron & Flaniken, 2011)

The method has several shortcomings because inaccuracies may arise where assumptions are made on the employee. Intentional and unintentional biases result in rating errors where the rater's judgement is influenced in one area by performance rating awarded in another area. Some evaluators may exercise extreme caution and consistently give average ratings to avoid being put to task on how the ratings were arrived at which punishes a performer while covering up for underachievement to the detriment of employee morale. Recent events can negatively influence scoring where a period of good performance can be dulled by a single unfavourable event occurring just before the evaluation is made .Another disadvantage of this type of appraisal is that it is highly subjective, has a focus on behavioural traits and has a limited range of application. For practicality, the method has to be used in conjunction with other techniques (Werner & DeSimone, 2011).

2.4 Police Performance Appraisal Procedures

The performance measurement systems of three modern policing agencies summarises best practice employed in those countries. The information was derived from interviews with police executives and other policing experts by the Rand Centre on Quality Policing, an institution that helps improve policy through research. This study focussed on the practices of the United Nations Police, and the Police Services in England, Wales and South Africa.

2.4.1 United Nations' Police Officers

The appraisal procedures for the United Nations' Police Officers are standardized into the United Nations' Accord, which are grounded upon the principles of professionalism, objectivity, transparency and impartiality in performance appraisal reporting. The reports shall highlight the positive as well as negative aspects of an Officer's conduct and competence as reflected in the performance of their duties and handling of their responsibilities (Chappell & Evans, 2008). The appraising Officers are made aware that in the process of reporting and reviewing performance that inaccurate assessments will reflect negatively on the appraisal of their own performance. Two key personnel are involved in appraising the performance of every UN Police Officer; the Reporting Officer who is the immediate line manager and the Reviewing Officer who is the Head of the Police Component or their delegate

Every officer upon attaining six months of continuous service with the peacekeeping mission must have a performance appraisal completed based on parameters stipulated in a standard prescribed appraisal form. Continuous service also includes periods of approved time off, annual leave, sick leave and any other certified period of absence from active duty. An officer qualifies to be rated as outstanding should they consistently perform above what is normally expected of them and shall not qualify based on a single exceptional performance. Performance benchmarks may be set by the Heads of Mission to reflect the unique operational requirements of a specific mission.

The police officer is required to peruse and comment upon the appraisal report before it is forwarded to the reviewing officer (Chapell & Evans, 2008). In the event the concerned officer is either unwilling or unable to comment on the appraisal, the reporting officer shall make appropriate remarks regarding the situation after which the form is submitted to the reviewing officer for final review. A reviewing officer will ensure that every officer is graded in line with the remarks made by the reporting officer.

In cases where an officer dissents with the ratings from the reporting officer, they may send a written counter statement to the head of the police component or their delegate along with a signed performance appraisal form within fifteen (15) days of the signature of the reviewing officer.

The head of the police component or his/her designate shall review these documents and shall take a decision to either overrule the counterarguments or call for a re-evaluation of the concerned officer (Chappell & Evans, 2008).

2.4.2 England and Wales Police Services

England has a very comprehensive system of measuring performance of all its law enforcement agencies that was instituted in 2004 (Rand, 2012). The assessment framework was developed to ensure quality in the performance measurement of police performance as compared to measures developed for other public services. These measures were designed to monitor progress towards achieving goals set out in the National Policing Plan.

An initial set of 35 measures was outlined to measure the performance of regional police forces. The statistics generated from audited data would be subjected to an annual comparison of performance across various police agencies. The Home Office developed a performance assessment website that would be available for scrutiny by the public who would make a quick assessment of the performance of their local police force relative to other police forces. Public scrutiny was therefore seen as a means of departing from a culture of focusing on measures based on crimes rather than those that affected specific local communities. The metrics further failed to address the disparity in the nature of crimes between the cities and rural areas and as such financial rewards favoured jurisdictions in which levels of crime were high.

The government then changed the performance measurement system by creating the police report card, which measured performance around four dimensions; local crime and policing, satisfaction and confidence, protection from serious harm and value for money. With these standardised metrics, a fair comparison could be made across all police services and jurisdictions. The publicly displayed results allowed citizens to examine their forces performance along the standardised metrics with an indication of whether the performance was improving, remained constant or whether the performance was on the decline. The website allowed comparison with that of similar forces. The only controversial aspect of this system was that the emphasis was on measuring public confidence.

2.4.3 South Africa Police Service

In South Africa, the performance management for police officers is based on a series of performance successes and errors. The successes police officers often present in their development of cases and arrest records can be a factor of several issues such as; the crime rate in the region and the affluence of the citizens.

The police reforms post-apartheid placed emphasis on policing as a service rather than a force as had been the case during the minority rule. Police oversight was seen as a major step towards realising much needed transparency within the various police services. To ensure effectiveness of the oversight mechanisms, mandatory performance reporting became necessary.

The South African Police Service recognised that police stations are the basic unit of policing service delivery and hence developed performance management charts that monitor individual police stations that combine 32 measures in four distinct areas of police services that are operational, information, resource and customer orientation. These indicators are monitored monthly, quarterly and annually.

The performance charts compare performance across various police stations and makes reference to a particular station's own performance within the previous reporting period. Targets at the beginning of every new reporting period are based on the station's performance over the previous four years with the most recent months bearing most weight. Numerical scores are calculated based on the percentage of the target score achieved. The performance ratings are made along five broad categories from excellent to unacceptable.

Performance data is simplified in a manner that allows police service staff to analyse at the provincial, station and incident levels to enable quick determination of relative improvement of various stations. Recent results point to an improvement in prosecution rates of priority crimes.

This system has however received criticism for heavily favouring the operational aspects of complaint investigation, emergency calls, offenses and bringing perpetrators to justice. Eighty-five percent of a station's overall score is computed from reported crimes, prosecution rates, response times and other operational indicators. This had undesirable effects such as achieving successful prosecutions through unorthodox means such as torture to yield confessions, which result in enhanced performance numbers. Citizen satisfaction is considered as part of the performance measurement index though it does not affect the overall performance score and remains unmeasured at most police stations (Faull, 2010).

2.5 Police Performance Appraisal Models

Studies undertaken in policing policy development has led to development of models along which police agencies intent on improving police performance and morale base their appraisal practices.

These agencies place emphasis on staff development and satisfaction of the employees as has been observed in Canada and Swaziland.

2.5.1 The Job Satisfaction-Based Model in Canada

According to Barrett, Peirone, and Cheung (2016), a survey representing 15 municipal Canadian police departments reported that performance appraisal systems in their departments fell short of established core features expected from a performance appraisal. Most officers cited that they hardly had the opportunity for input into their appraisal, they did not receive regular feedback and the appraisal ratings were based on personal traits rather than performance metrics. They also indicated that the appraisal results did not lead to improvement in performance of duty or training needs assessment as expected.

The survey also revealed that supervisors received little or no training as far as conducting performance appraisals was concerned. This led to development of a model whose outcome was to ensure the police officers were satisfied with their jobs as depicted below:

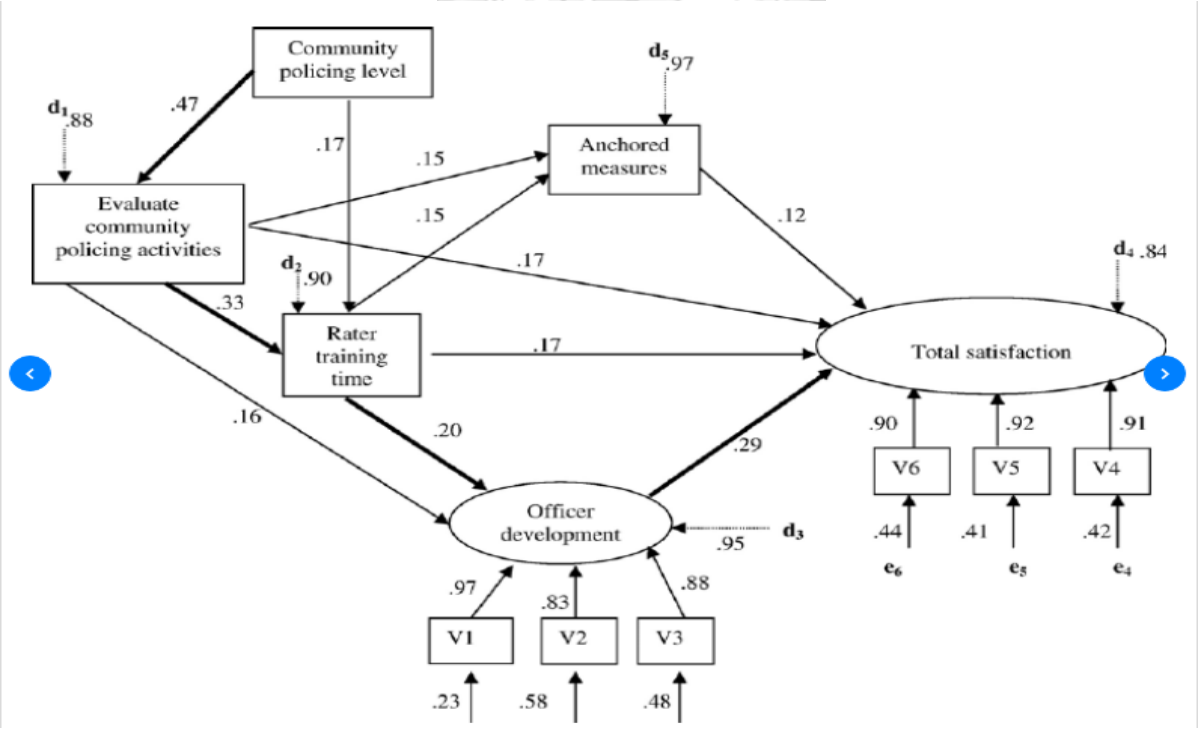

Figure 2-1 The appraisal model based on job satisfaction (Barrett, Peirone, & Cheung, 2016)

The officers involved in the survey indicated that feedback on their performance was the most desired outcome of the appraisal process.

2.5.2 The Reward-Based Model in Swaziland

Mabila (2014) proposed a reward based model aimed at motivating members of the Royal Swaziland Police Service (RSPS). She based this model on the premise that organizations need to provide the environment, support and resources that impact favourably on the motivation of individuals. Motivation comes in the form of rewards, learning opportunities and encouragement that maximize the chances that employees will be satisfied and therefore perform optimally. There exists a relationship between job satisfaction and job performance and there is sufficient evidence to indicate that job performance results in job satisfaction (Chung et al., 2017).

Mabila's research in (Mabila, 2014) found that police officers are more likely to perform effectively when given necessary resources. The findings also showed that police officers work for long hours and their efforts are not adequately rewarded, they are not treated fairly and there are insufficient structures to help police officers who experience burnout. In exploring the impact of job satisfaction on performance, the study established that job satisfaction and work performance cannot be separated. The proposed model by Mabila (2014) is as shown in Figure 2.2.

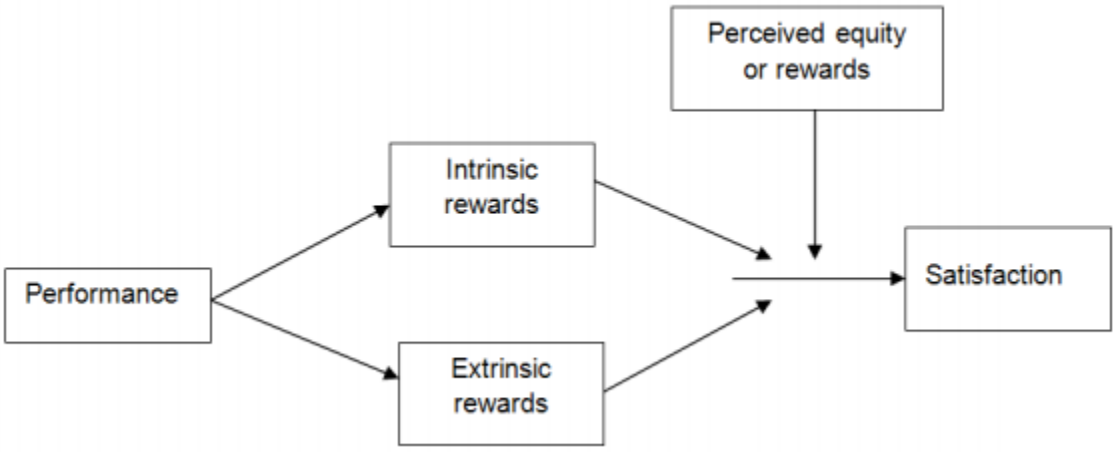

Figure 2-2 The Reward-based model (Mabila, 2014)

2.6 Algorithms and Models Used in Performance Appraisal Systems

Algorithms often provide for a model-based approach to handling the appraisal of employees. Employee suitability and effectiveness is determined through performance appraisal in most organisations. The performance of a department has been found to be dependent on a number of parameters broadly ranging from the individual's qualifications, experience, level of commitment, research activities undertaken towards institutional support, financial feasibility, and top management's support. These parameters act as performance indicators for an individual and group and subsequently can affect the decision making of the individual and the stakeholders. Algorithms combine all these sources of appraisal decisions into a singular model that can then be applied effectively in the appraisal process.

2.6.1 The PageRank Algorithm

PageRank (PR) is an algorithm used by Google Search to rank websites in their search engine results with an underlying principle of assigning a certain numerical value to every website (Chitroda & Katurbadham, 2014). This numerical value is determined by the number of links that point to a particular website. Links emanating from highly ranked websites are considered to be of greater value. The smaller the number of links on a particular website leading to the target website then the higher the quality of each of the links. It was used by Google to rate the performance of its employees in the early 2010s. According to Kwon, et al. (2012), the algorithm presented a criterion for determining how best to assess that an employee was performing above par or that their numbers were good enough. Kwon, et al. (2012) intimates that Google also sourced the algorithm to other companies in Silicon Valley to evaluate the performance of employees who were not always under strict scrutiny. These employees would hence be considered to have worked successfully in roles that were greater than the said job specifications as required by Google.

The PageRank of the aspect being investigated can be considered as the best possible result compared to a set of results. The algorithm is thus recursive and depends on the number of comparisons to come up with the most efficient result, given a set of deliverables and aspects that need to be compared. The PageRank can be considered to be the view of the page that represents the most desired information among a set of pages (Pälli & Lehtinen, 2014). Similarly, in using the algorithm to determine the performance of an individual, it may suffice to have an aspect of quality or excellence in mind. It is

such quality that guides the overall investigation process, to ensure that the desired outcome is eventually achieved.

The PageRank algorithm is effective where the chosen appraisal technique is that of paired comparison. This technique consists of comparing an individual employee to another or one team to another in areas such as efficiency with clear metrics such as meeting deadlines and economy in the use of resources (Kaminska & Szymanczyk, 2014). An indicator is then originated and is eventually used in ranking. Use of this algorithm in determining appraisal scores is based upon assumption that the appraisal may not be one-sided, time consuming or incomprehensible to the personnel who take part in it. In this way, preferences of the staff can be taken into account especially in small organisations where colleagues are well familiar to each other or within departments in the case of large firms.

The employees are asked to fill in a short questionnaire in which they provide basic information pertaining to themselves and state who among their colleagues makes the greatest impact, in their opinion, to the organisation or department. This is where the association with the PageRank algorithm is brought out by the assumption that each employee is assumed to be running a website and the selected colleagues are recommended links. This algorithm is suitable for such ranking because good employees are able to judge who else is good at performing tasks that they themselves perform well at. The more the number of staff nominated by one person then the lower the value of that ranking since this is an indication of indecisiveness. The ranking therefore takes into account both the number of votes and their origin (Kaminska & Szymanczyk, 2014).

At this point, an employee whose name features in a questionnaire has been rated highly and those that do not appear have been rated low meaning the end result is information regarding the general virtues of employees. To address this flaw, more information can be requested from the staff for example introducing a rating scale of 1 to 5 with 5 representing a high rating and 1 the lowest rating giving the appraiser the opportunity to nominate as many colleagues as one desires within each rating scale. Appropriate weights can then be introduced to ensure diversification in appraisal without having a questionnaire that is complicated (Kaminska & Szymanczyk, 2014).

The PageRank algorithm is practical in determining ratings amongst superiors, peers and subordinates whose interaction is essential in achieving organisational goals.

2.6.2 Cumulated Average Weighting

The cumulative average weighting method is a mathematical model that was developed to make determination of an individual's appraisal performance score more representative and unbiased (Ahsan, 2016). The method considers a set of predefined variables each of which is evaluated by weighted values of the corresponding set of attributes on a Likert scale.

The formula for score calculation is based on score determining variables, attributes, mathematical operations (cumulative average and arithmetic mean) and the reflective factor. The reflective factor refers to the opportunity factor and the optimum contribution factor. The opportunity factor relates to individual expertise and relevance of skills to the particular job description. This factor ensures uniformity in assessment for all concerned employees. The optimum contribution factor takes into account the individual's overall contribution over their entire period of service.

The Cumulated Average weight is a representation of the average scores given to an individual assessment, from different sources. The cumulating score can also be represented in the form of the variance of the score, to offer a more realistic look into the performance record of the activity, as indicated by other relevant parties. The outcome is that the bias of any of the reviewing individuals is spread across a spectrum and highly reduced by the time the final score is offered. The benefit of the approach is that; the overall score is a true indication of the different modular units involved in the appraisal process (Ahsan, 2016).

The naïve cumulated average weighting algorithm can be used to get the variance in a set of appraisal scores for an individual (Ahsan, 2016). The simplistic pseudo code for this algorithm is as shown in Figure 2.1.

- **Let ← 0, SumSq ← 0**
- **For each datum x:**
  - **n ← n + 1**
  - **Sum ← Sum + x**
  - **SumSq ← SumSq + x *x**
- **Var = (SumSq – (Sum * Sum)/n)/(n – 1)**

Equation 2.1: Pseudocode for Naïve CAW Algorithm

The equation behind the Naïve CAW is as shown in Eqn. 2.2.

$$\sigma^2 = (\overline{x^2}) - \overline{x}^2 = \frac{\sum_{i=0}^{N} x_i^2 - (\sum_{i=1}^{N} x_i)2 \ /N}{N}$$

Equation 2.2: The Naïve CAW Algorithm

The two-pass cumulative average weighting algorithm first computes the sample means of the appraisal results from a Likert scale then computes the squares of the difference of the means to get the standard deviation. From the standard deviation, the average appraisal score can be determined from a set of unarranged data elements. This is as shown in Equations 2.3 and 2.4:

$$\overline{x} = \frac{\sum_{j=1}^{n} x_j}{n}$$

Equation 2.3: Computation of Sample Means

$$variance = s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}$$

Equation 2.4: The Variance Computation

The pseudocode for the two-pass CAW is as shown in Equation 2.5

```
def two_pass_variance(data):
        n = sum1 = sum2 = 0
        for x in data:
                n + = 1
                sum1 += x
        mean = sum1/n
        for x in data:
                sum2 += (x – mean) * (x – mean)
        variance = sum2/(n - 1)
        return variance
```

Equation 2.5: Pseudocode for implementing the two-pass CAW algorithm

The computation of the average score can take any form of the averaging modules represented in the different data structures available in different language libraries. Essentially, it is not important to determine which order of summation the algorithm takes. The most vital aspect of the algorithm is to ensure that the benefit of the user in acquiring a measure of central tendency is gained, from assessing measures of dispersion. Data about varied issues in the performance report is aggregated using arithmetic equivalence to come up with a figure that represents the proportionate value of the entire research. The exercise can thus only be considered to be valuable and reasonably consistent with the diversified sources of appraisal data that can come from alternating sources yet significantly affect the appraisal score of the individual.

2.6.3 Fuzzy Based Multifactorial Evaluation

A performance appraisal is typically a multiple criteria decision-making process, which according to research (Meenakshi, 2012) can be successfully solved by applying fuzzy set theory. The 360-degree evaluation method has been found to be comprehensive since the ratings pertaining to a particular employee comes from points of contact as they discharge their duties i.e. peers, managers, subordinates, internal and external customers and suppliers leading to more transparent and beneficial results compared to other appraisal procedures (Meenakshi, 2012).

To maintain the usefulness and relevance of an appraisal system to management, an appraisal system must consistently produce reliable results while taking into account deficiencies given the nature of human behaviour. Fuzzy based methods and electronic nominal group technology have been used to solve multi criteria problems that include multi-criteria performance evaluation in various organisations. In many cases, appraisers have to rely on vaguely defined qualitative criteria in evaluating the performance of their subordinates, which creates difficulty in consistently quantifying the score for each individual. Measuring performance and in particular improvement in performance has a certain level of ambiguity and fuzziness (Shah et al., 2014).

The application of fuzzy set theory on computer-based fuzzy group decision support systems and in particular the multifactorial evaluation model has been found to assist decision makers in making better decisions by managing vagueness inherent in performance appraisals. The appraisal process involves evaluation of an employee based varied criteria which makes it a case of multi criteria decision-making problem that can be successfully solved by fuzziness (Ozkan et al., 2014)

The values for the qualitative assessment criteria cannot be precisely defined, quantified and replicated by decision makers. It is not easy to quantify ratings leading to fuzziness and therefore fuzzy decision making tools can solve the problem of ranking fuzzy numbers. A fuzzy group decision support system would therefore increase appraisal productivity while making the appraisal result transparent and logical (Meenakshi, 2012). The fuzzy set theory is ideal in developing performance appraisal systems because it allows use of fuzzy variables and relationships to effectively and consistently produce accurate appraisal results and can be adjusted with ease whenever necessary.

The method enables grouping of employees with respect to certain criteria that enables implementation of informed human capital management decisions such as training, promotion or salary increment (Ozkan et al., 2014). Machine-learning algorithms strengthen the weaknesses of conventional methods of appraisal since they enable grouping of employees with respect to several clusters e.g. those deserving promotion, those due for a pay rise, those in need of training etc. Conventional rating method assign membership of an employee strictly to a particular cluster whereas in the case of machine-learning, membership degrees to a particular cluster are computed in a manner that allows easy decision making in the event an employee is placed within more than a single cluster conventionally (Yee & Chen, 2009).

The use of fuzzy logic in performance appraisal allows inputting evaluation parameters in the form of linguistic variables that do not have distinct quantifying values but have values lying in between a certain range. The system allows the decision maker to introduce vagueness and subjectivity, which mimics human-like behaviour in decision-making. It creates a system of appraisal, which is consistently able to produce reliable and valid results for the appraisal process (Shah & Ladhake, 2014).

2.7 Machine Learning Algorithms in Performance Management

Machine learning (ML) is the subfield of Artificial Intelligence (AI) that gives computers the ability to learn and act like humans, without being explicitly programmed, by feeding data and information. The system is trained by the past data to build a prediction model, enabling it to recognise complex patterns and to make accurate decisions based on the training data set that tells right from wrong until it has enough answers to predict an accurate as possible output (Alzubi et al, 2018).

Machine Learning is concerned with the design and development of algorithms, which are organized, based on expected outcomes. The machine learning algorithms learn and optimise their operations by analysing input data to make predictions within an acceptable range. With the feeding of new data, the algorithms tend to make predictions that are more accurate. Machine learning techniques have the ability to generate intelligent decisions given the huge amounts of data available combined with increasing computing power by implementing the Knowledge Discovery in Database (KDD) approach (Mahesh, 2020).

Machine learning has been used in the design of efficient performance evaluation systems void of drawbacks exhibited by conventional appraisal methods. Model development takes place in several steps. The first step is gathering data, since sufficient historical data is required to train the model. This data is then pre-processed since the system cannot use it while it is still in its raw form. Then an appropriate model is chosen which is then trained to convert a blank model into a trained model. The model is then tested to make sure it is working fine with minimal errors. It is then routinely fine-tuned for increased prediction accuracy. The model is then deployed to perform predictions.

Some machine learning algorithms that have been used in evaluating employee performance include Decision Tree, Random Forest, Naïve Bayes, K-Means and Support Vector Machine.

2.7.1 Decision Tree Algorithm

Decision Tree is a supervised machine-learning algorithm and is one of the earliest and most widely used machine learning algorithms. The model is represented by an upside down tree-like structure with the data set being broken down into smaller and smaller subsets as the tree grow. Each node represents a feature (attribute), each branch represents a decision and each leaf (terminal node) represents an outcome (categorical or continuous value).

The very top of the tree is called the root. Then there are internal nodes or branches and they have arrows pointing away from them. The last levels are called leaf nodes or just leaves. They have arrows pointing to them but do not have arrows pointing away from them.

An internal node or decision node contains a simple Yes/No question to be answered about a particular variable. Based on this the tree splits into branches, one branch representing “Yes” and the other “No”.

At the end of each branch is another question. This recursive process continues and the tree gets deeper and wider with each decision made until it terminates in a leaf node that makes a prediction.

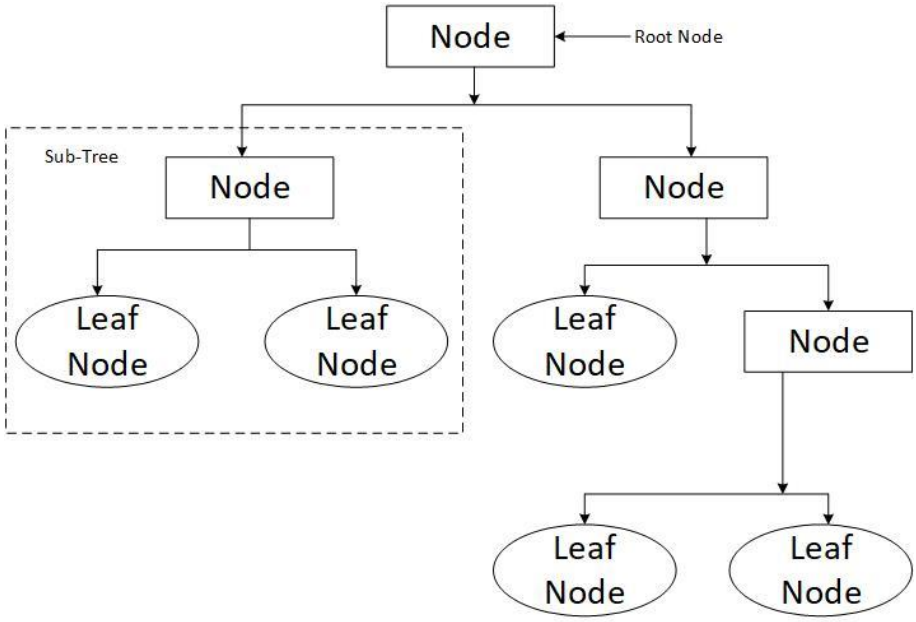

Figure 2-3 Decision Tree

Decision Trees produce human readable classification rules enabling easier interpretation compared to other classification methods. It is therefore ideal in predicting the performance and the future development needs of an employee from a set of training data (Magesh et al, 2013). The challenge in implementing decision trees is that they are sensitive to noise that requires modification of the training set that in turn will change the decision tree.

2.7.2 Random Forest Algorithm

Random Forest is an ensemble classifier based on decision trees used to improve accuracy by creating multiple decision trees using the dataset of original data and randomly selecting a subset of variables at each stage of the decision tree. Single Decision Trees are sensitive to their training data which makes them prone to error from the test data. When using the Random Forest approach, the individual Decision Trees are trained using different parts of the training data set. To classify an unknown sample, the data is passed down each Decision Tree in the forest. The model selects the mode of all the predictions of each decision tree which reduces the risk of error from individual trees.

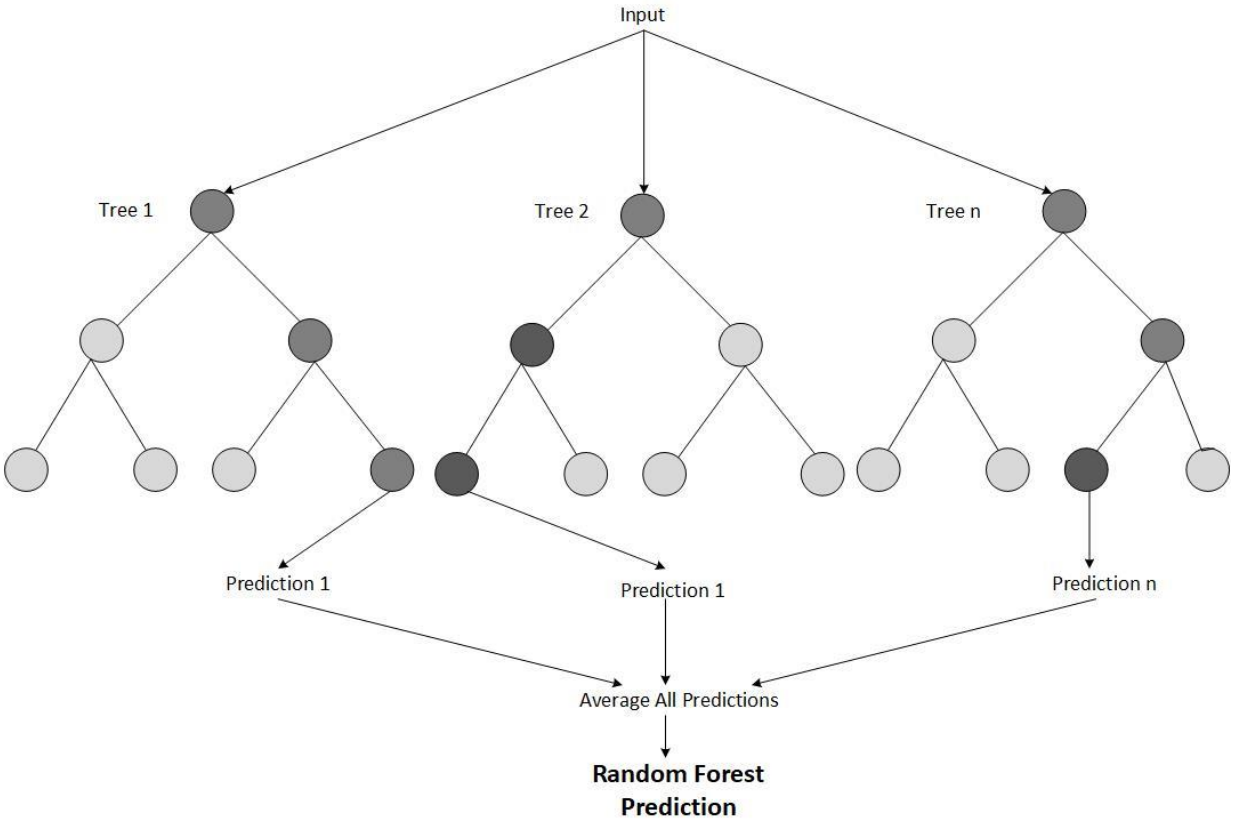

Figure 2-4 Random Forest Structure

This technique has been used to monitor the performance of employees to monitor variation in performance and predict employee turnover, which is a major problem for many organisations. The result of this prediction helps organisations to either plan for this turnover or implement employee retention measures especially in key positions (Gao et al, 2019).

2.7.3 K-Means Clustering

The K-means is an unsupervised learning clustering technique is a very efficient algorithm used in data mining. The algorithm organises unlabelled data by creating partitions within a data space in k-number of clusters through a machine learning process. These clusters contain data points that are similar to each other and dissimilar to data points contained in other clusters. The algorithm achieves this in steps by first deciding the number of clusters to be identified within a data set, which will be the starting value of k. These initial clusters are randomly selected from distinct data points. Assuming

that the data is linear, the distance between the data points and each of the initial identified clusters is measured (Sinanga & Yang, 2020).

The next step is assigning the remaining data points to the nearest initial point with a k value. The assignment can be made to one or more of the initial values. This step is using the mean distance values between the k data points and the other data points Where the data is non-linear, the clustering process still works by initial data point selection the taking the Euclidian distance between points using the Pythagorean theorem then calculate the centre point and re-cluster. The algorithm settles on a value of k when the assignments no longer change (Sinanga & Yang, 2020).

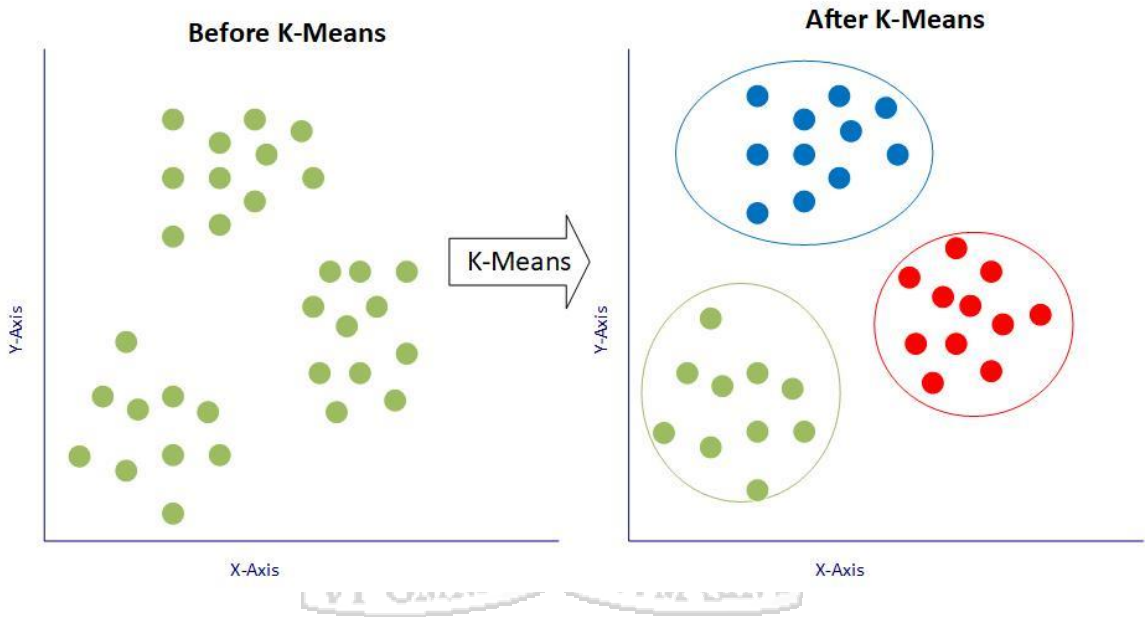

Figure 2-5 K-Means Clustering

This algorithm has been used to automate human resource needs deployment needs due to its simplicity in implementation and efficiency. It has been used for both initial placement and in predicting future departmental distribution based on such factors as risk handling, managerial acumen, current compensation rates and performance metrics (Kakulapati et al, 2020). Through clustering weakness in deployment in deployment are identified along with potential solutions. These analytics have enabled human resource decision making that improve employee satisfaction and increase innovation.

2.7.4 Support Vector Machine

Support Vector Machine (SVM) is a powerful supervised machine learning technique that has a simple structure and good classification ability with small data sets. SVM works by first separating input data by drawing a linear boundary between (the decision boundary) them, the maximum margin hyperplane (Nguyen, 2017). Several boundaries can be drawn therefore a suitable solution would be to have the boundary as far as possible from any of the data points giving a maximum margin. The wider the margin between the positive and negative hyperplane the more accurate the classifier. The optimal solution is determined only by the data points nearest to the margin, which are the Support Vectors as shown in Figure 2.3 below. These are the data points that are most difficult to classify. Moving the support vectors has an impact on the decision boundary whereas the other data points have no impact on the decision boundary. In this way, the algorithm gives the best possible separation between classes without misclassifications (Bridgelall, 2017).

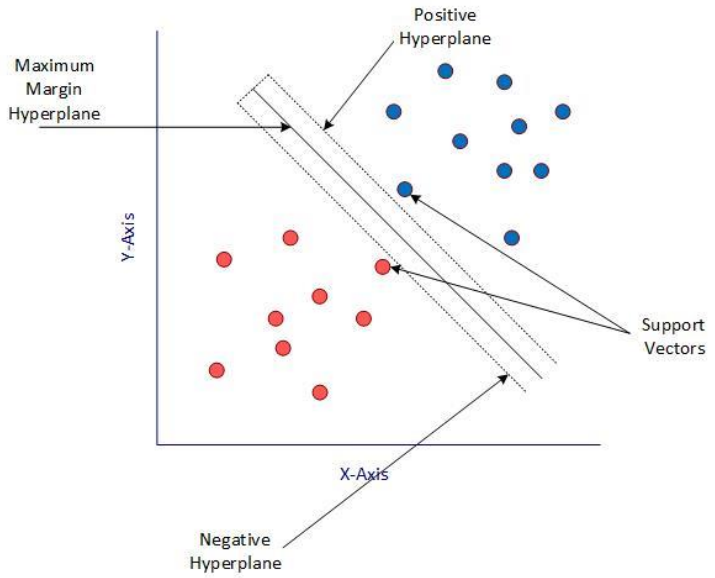

Figure 2-6 Classification of Data by Support Vector Machine Algorithm

In human resource management, SVM has been used to manage talent management and employee selection due to its ability to solve complex optimisation problems. The technique has been found effective in identifying the required data pattern for classifying an employee's achievement, which in turn may form a basis for promotion (Jantan et al, 2014). It can also assist in training needs analysis

and matching employee skills, knowledge and personality to determine their right placement within the organisation.

2.7.5 Naïve Bayes Classifier

This is a supervised learning classification algorithm based on the Bayes Theorem and acts like a probabilistic machine-learning model used for classification of tasks. The Bayes theorem is an important theory in probability and very useful in machine learning. It can describe the probability of an event based on the prior knowledge of conditions related to that event and is represented as follows:

$$\mathbf{P(A|B) = P(B|A)\,P(A) / P(B)}$$

Equation 2.6 Bayes Theorem

Bayes Theorem gives the conditional probability of an event A taking place given that another event B has already taken place.

Using this algorithm, it is fast and easy to predict a class of test data set and is very useful for large data sets where it outperforms other classification methods. The learning model is computed from a set of conditional independences within the training data However, the algorithm is unable to make a prediction when an input value is not present in the training data set since it would return a zero value upon multiplication. This algorithm is mainly used for clustering and classification and is dependent on conditional probability which assumes that attributes do not depend on each other (Himanshu et al, 2020).

The Naïve Bayes classification algorithm has algorithm has been used to predict the performance of employees. It takes into account every single attribute regarding employee that contributes to their performance and can be used to make critical management decisions such reassignment (Jayadi et al, 2019). For accurate prediction however, this classifier requires a lot of training data.

2.8 Research Gap

There are very few tools that can be used to ascertain the effort the police officers in Kenya make (Makabira & Waiganjo, 2014). From the surveyed literature, best practice in appraising a police officer requires input from the officer, a customer satisfaction perspective, fair supervisor assessment, professional opinion of a peer and a system whose output will address the development needs of the

officer. The proposed model will fill the gap in the Kenyan case by introducing quantitative aspects to the current appraisal system, which is largely qualitative.

The concept of multisource feedback that eliminates bias that occurs when final scores are determined by a single source and also serves as a developmental tool that aims to improve the performance of an officer from the perspective of all who interact with them as they perform their duties. It thus provides an objective platform to measure performance of individual officers.

The proposed model will take on a hybrid approach where the Cumulative Average Weighting algorithm will aggregate scores as reported by the members of the public (customer feedback). The 360-degree appraisal technique is a Multi Criteria Decision Making (MCDM) procedure, which is subject to ambiguity that leads to bias. Fuzzy logic is one of the major platforms used to deal with vagueness. The use of fuzzy logic is expected to make performance scoring unbiased by taking into account reflective factors with several variables to eliminate falsification of reports in the reviewed appraisal systems.

2.9 Fuzzy Logic

Fuzzy logic was chosen in this study to develop a model that would determine the performance score given multiple appraisal score sources. This selection has been informed by the need to develop an appraisal model that will introduce objectivity in ratings through rules rather than mathematical expressions. The performance evaluation would be better defined across a continuous scale rather than by discrete outputs hence the suitability of fuzzy logic, which resembles human reasoning (Suh & Kim, 2020).

Fuzzy logic is most suitable in describing linguistic variables qualitatively (Gallab et al., 2019) since some events cannot be described numerically. It effectively deals with subjective, incomplete or unreliable knowledge bases. Fuzzy sets were introduced to represent the values of real world parameters whose boundaries are not crisp due to the subjectivity of human measure and ambiguity (Siraj, 2019). A fuzzy logic system is developed with three major modules; the fuzzifier that will generate fuzzy output membership function graphs, an inference engine that in this study will run off IF-THEN rules and a defuzzifier that will generate crisp outputs using the centroid method.

2.9.1 The Fuzzy Inference Engine

The fuzzy inference engine will map the four inputs to a single output based on a set of pre-defined rules in the rules base. Use of a fuzzy rules base provides a practical alternative to mathematical functions. Mathematical models provide good theoretical academic models that are not easily applicable in solving real world problems. The effectiveness of a fuzzy inference engine is determined by the accuracy of the rule base. A model with more than two inputs presents an additional challenge in that the number of rules will greatly increase for each added input. Use of a fuzzy associative matrix table (FAM) simplifies the rules generation process and gives a visual to the person developing the rules (Siraj, 2019). Below is a typical structure of a fuzzy inference engine.

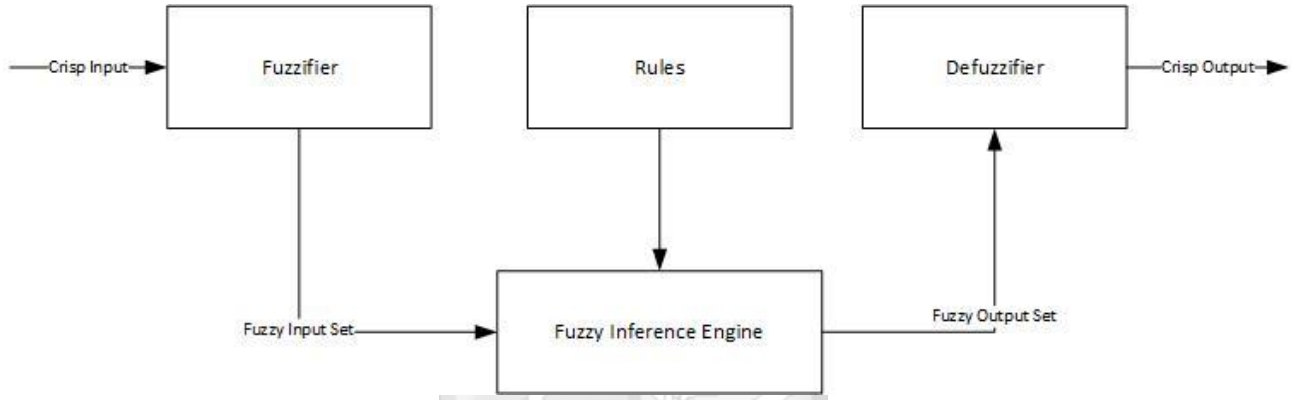

Figure 2-7 Fuzzy Inference System

The modules of a fuzzy inference system function as follows:

i. **The Fuzzifier** – this module receives crisp input values from a proposed fuzzy set and converts it to fuzzy subsets. A membership graph with values between 0 and 1 maps these inputs to two corresponding values where applicable. The rule of maximums (OR) or minimums (AND) is used to select the most appropriate value.

ii. **The Rule Base** – this is where a set of fuzzy rules that are required to initiate the model's logic are stored. These rules are usually in the form if IF-THEN statements developed either from intuition by experts in the field of application or from data as is the case with self-learning algorithms.

iii. **The Inference Engine** – this module infers knowledge from fuzzy inputs using the IF-THEN rules by assigning to the corresponding output as per the definitions.

iv. **The Defuzzifier** – this module converts the fuzzy output from the inference engine into crisp values using techniques such as the Centroid method.

2.10 Conceptual Framework

The dependent variable that is output of the system process is the Multi-source Feedback that will be the officer's Performance Score. The output translates to an officer's performance as either exceptional, superior, satisfactory or below average. The independent variables are the ratings given to each officer by the public, supervisors, peers and subordinates that form crisp inputs to the Fuzzy Inference System.

The sequence to generate crisp input from the public will be initiated when a complaint is raised to the Police via the system's mobile front end and based on the selected category, the system will assign an appropriate officer to handle the reported complaint. The attending officer logs a report back into the system once the assignment has been completed, a copy of which is shared with the complainant.

The system will then request feedback from the complainant on their experience from the time they made the complaint to the officers' responsiveness, the accuracy in reporting and general handling of the incident. The response will be based on a scale, the more favourable the response from the complainant the higher the score. The scores from individual ratings are then combined to form an overall public score that is aggregated for the entire period an officer has served in various stations.

This score along with ratings from the immediate supervisor, peers and subordinates where applicable are fed into the Fuzzy Inference System as crisp inputs. The overall performance score, which is a measure of an officer's effectiveness, is determined after application of fuzzy rules and defuzzification.

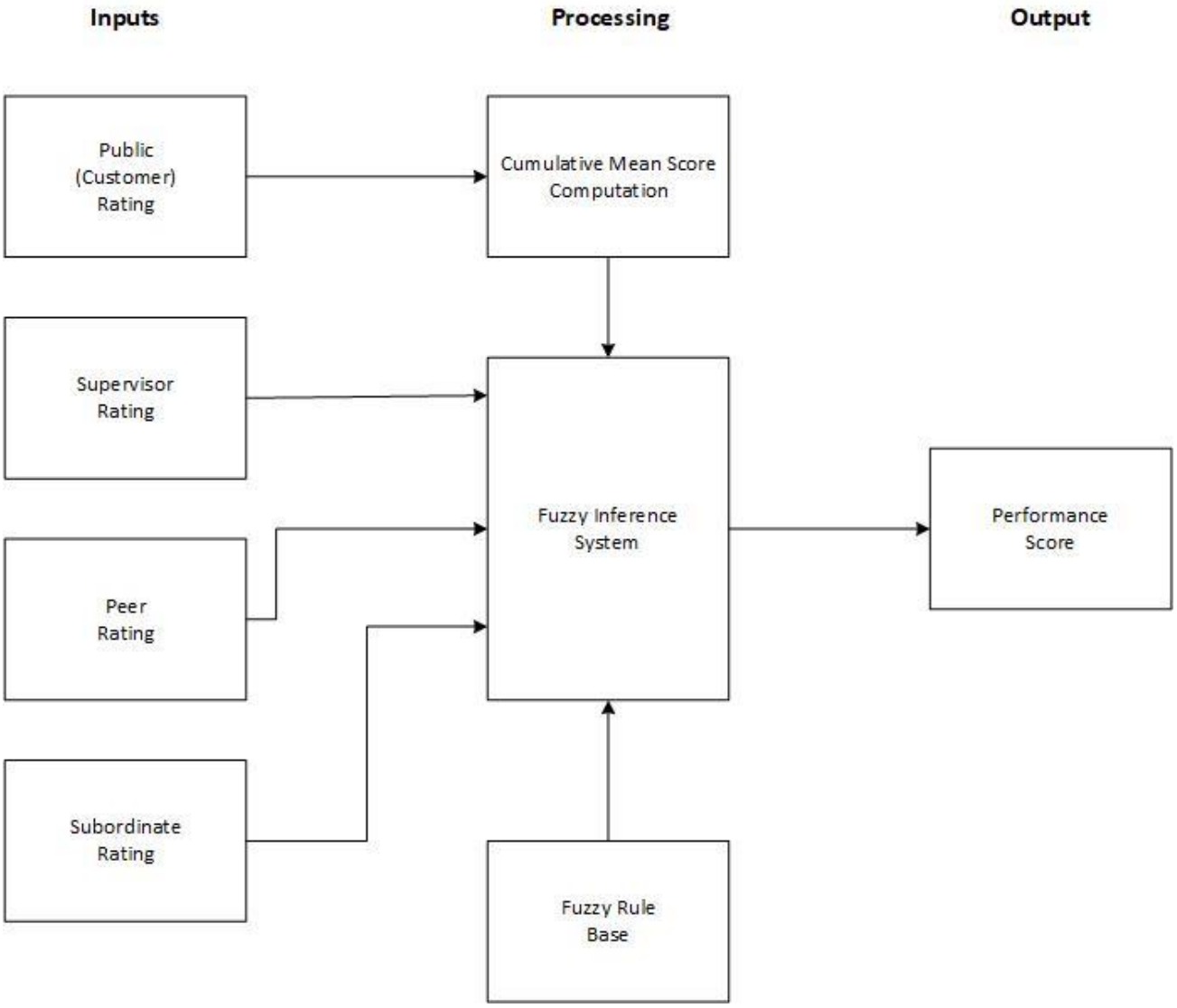

Figure 2-8 Conceptual Framework

Chapter 3: Research Methodology

3.1 Introduction

The research methodology lays out the researcher's strategy and facilitates resource efficient research (Saunders, Lewis & Thornhill, 2013). The correct methodology would therefore ensure that the resultant information system would solve the identified problem.

This chapter presents the methodology used, the type of data required to run the model, data collection, analysis, and the model development tools that were used. The study's reliability and validity were discussed as well as ethical consideration while undertaking the study.

3.2 Research Design

Research design constitutes the blueprint that aids the researcher in the allocation of limited resources by posing crucial choices in methodology (Saunders et al., 2013). Qualitative research was used to gain an understanding of the performance appraisal process of law enforcement officers and identify what would qualify as best practice as well as the shortfalls in current systems.

This study employed an exploratory research design by examining relevant data in regards to performance appraisal. Exploratory research enables a researcher to expand their knowledge of the subject in the shortest time possible and allows for efficient resource utilisation through narrowing in to material relevant to the study objectives. The acquired knowledge could then be used to provide a practical solution to the stated problem.

Theories and procedures from the existing body of knowledge in regards to employee performance appraisal were explored, which informed the design and development of a multi-source feedback appraisal model for police officers. The conceptual framework was used as reference in developing the preliminary system design and thereafter the functional and non-functional requirements informed the detailed design. The design was implemented by independently building the constituent modules of the 3-tier architecture, verifying their functionality then integrating and testing the prototype.

3.3 Model Development

The Fuzzy Expert model was developed through several phases. The implementation was undertaken through four distinct steps.

3.3.1 Model Data Generation

The purpose of fuzzy logic in this study was to enhance objectivity in an appraisal process that left to human reasoning and judgement would be prone to subjectivity. The data that built the Fuzzy Inference System was derived from officer ratings from all the multi-rater variables; the customer (public), supervisor, peers and subordinates where applicable depending on the individual's rank. All officers must have the customer, supervisor and peer components. The output data – the performance score will be the fuzzy set with fuzzy subsets indicative of an individual's general rating as determined by the fuzzy rules.

3.3.2 Customer Rating Generation

This component of the FIS crisp input data was derived from the mobile application. The user initiates the model sequence by reporting an incident and completes their interaction with the model once their rating based on pre-defined criteria. The mobile application provides features for user registration, reporting, attending officer input and user experience rating. This dataset is a key input to the FIS and the individual officer's ratings are aggregated over time taking into account all ratings awarded by various users. To ensure uniformity in score computation and aggregation, the cumulative average-weighting algorithm was used.

3.3.3 Supervisor, Peer and Sub-ordinate Input Data

Officers having the rank of Corporal and above will have components of their supervisor, peers and sub-ordinates ratings. Police Constables, having the lowest rank, will not have a sub-ordinate rating component. This crisp input data was obtained from the web application through which the respective parties concerned record annual ratings. The web application contains features that enable ratings to be performed at pre-determined period within the appraisal period.

3.3.4 Fuzzy Rule Development

The fuzzy rule base is core to the functioning of the fuzzy logic model. The rule base requires input data in the form of statements that inform assignment of a particular output score within the fuzzy subset. A user familiar with the expected real world output will develop these fuzzy rules.

3.4 System Development Methodology

System development is a continuous and formalized process given the dynamics of the technological landscape. The proposed module was developed following gaps identified in surveyed literature. The intended users interacted with the prototype to understand what the system module would achieve and how best to implement it. System prototyping was therefore the most suitable methodology due to its flexibility that allows alteration at any point in the development life cycle. The prototyping methodology is as illustrated below:

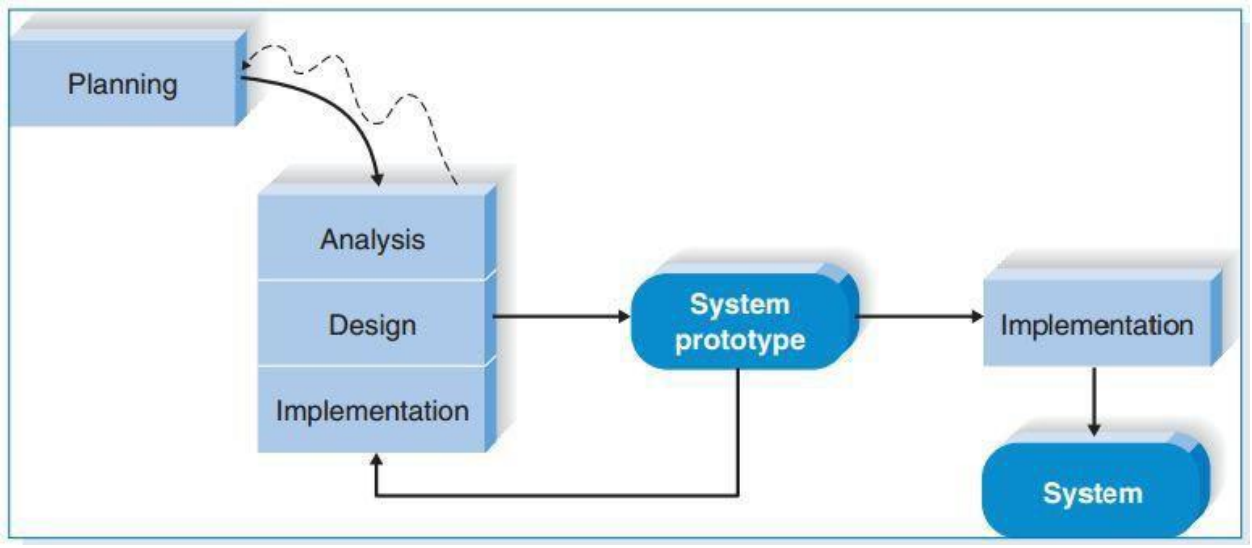

Figure 3-1 System Prototyping (Dennis, Wixom & Roth, 2012)

According to Hoffer (2012), system prototyping works by performing the analysis, design, and implementation phases concurrently in order to quickly develop a simplified version of the proposed system and give it to the users for evaluation and feedback. Since the initial version provides minimal features, this cycle is repeated to correct undesirable features and for addition of more features.

3.4.1 Requirements Planning

During the planning phase, the proposed application's functional and non-functional requirements were determined. These requirements were determined by defining the problem and reviewing relevant literature to establish gaps that justify gaps that can be filled by the proposed application and solve the identified problem. Process observation in the course of the researcher's duty was used to understand the challenges with the current system. The system development work-plan was also prepared during this phase.

3.4.2 System Analysis

Use case modelling was done to enable identification of the system actors and create an understanding of the system's functional requirements. The sequence of execution of the interrelated activities was outlined with clear understanding of what activity within the proposed system each actor initiates. UML diagrams were used to define use cases, data flow and activity sequence which guided the implementation through coding.

3.4.3 System Design

The design phase focussed on fulfilling the proposed system requirements defined during the analysis phase. The gathered requirements were used to determine interaction of the various sub-systems including implementation of data storage. Entity relationship diagrams were used to visualise the relationship between objects, entities and activities within the relational database. The system inputs and outputs were mapped and implemented as per every actor's role within the sequence of activities.

3.4.4 System Implementation

The proposed system was actualised during this phase through coding and testing the prototype. The mobile front end was developed using the Android SDK, Java was used to build the web based server side and the application layer of the system. MySQL was used for the relational database development through which access controls for the various levels of users were defined. Matlab was used to implement the Fuzzy Inference System from which the performance score would be determined.

3.4.5 System Testing

Acceptance testing was carried out on the prototype to determine whether it met the desired requirements and ensure that it solves the research problem. Bugs in the code were identified and

resolved. Load testing was carried out by sending simultaneous requests by different users to determine the number of requests the application would handle at peak usage. User acceptance testing was undertaken through feedback from users.

3.5 Location of the Study

The location for this research was Langata in Nairobi County. The target population was identified during the testing phase to determine usability of the prototype and to inform necessary adjustments.

3.6 Research Quality

The quality of this research was assessed based on the validity and reliability. The prototype was tested to ensure that computation of appraisal scores was consistent based on the ratings awarded by the various users.

3.6.1 Reliability

Reliability refers to the extent to which results are consistent over time and that they can be reproduced at different times and under different conditions. Re-tests were done to measure consistency of results of data input and system generated scores.

3.6.2 Validity

Validation was carried by testing the prototype to ensure accuracy of the algorithm in determining scores and to minimize input of wrong data. This ensured that the application met its intended purpose.

3.7 Ethical Considerations

Ethics in research ensure that no one suffers adverse consequences from the research activities that require personal integrity on the part of the researcher (Saunders et al., 2013). The researcher ensured that the study benefits were explained to all respondents involved in the testing phase of the model who were informed that their involvement was strictly for purposes of this study. The researcher cited work from secondary sources giving due credit to the authors.

Chapter 4: System Design and Architecture

4.1 Introduction

This chapter evaluates the conceptual and logical aspects of the proposed prototype by explaining the design aspects of the various modules. A detailed understanding of the system design was illustrated through UML diagrams (Use Case, Sequence, Entity Relationship, Data Flow and Class diagrams). These diagrams define the system modules and components.

4.2 Requirements Analysis

System requirements gathering precedes system analysis. The requirements gathering process involved reviewing literature that gave insight into systems in use in police appraisal. The literature provided insight into appraisal parameters that can be included into an information system. The functional and non-functional requirements of the appraisal prototype were identified in the following section.

4.2.1 Functional Requirements

Functional requirements define what the system is intended to perform. They include specific functions that describe how each use case will be executed. The functional requirements were derived from an analysis of secondary information sources on similar appraisal tools through which standard functionality these systems were identified. The use of secondary sources of information helped in identifying the approaches used in appraising the performance of police officers and the gaps within these systems. Gaps in these systems informed improvement as well introduction of new functionality. Additionally, the actions that each actor would be expected to execute within the system were considered in determining these requirements.

The model will have a graphical user interface that will form the access point to the system. The interface shall allow connection to the system dashboard, taking input from the user to report crime incidents and rate officers as well.

The system shall generate officer-rating scores based on scores from the customer's feedback as well as the supervising officers. The scores shall be retrieved upon request by an authorised officer and will at a glance present an average score as well as detailed performance trends across service stations. The

officers as well as their supervisors will be granted view only access to these performance trends. The system should be able to execute the following:

i. **User management** – the administrator will be able to administer the system and can add or delete users, enable or disable users and define rights for all users as well as assign administrator pledges.
ii. **User registration** – members of the public will be able to register a user account in the system.
iii. **Log in and log out** – all registered users of both the mobile and web-based clients should be able to login to access system functions and log to terminate the session. Automatic log out will be executed when the system is judged to be idle.
iv. **Incident reporting** – members of the public will be able to report incidents and include brief relevant details. The reporter's location feature should be on at the time of submitting the report.
v. **Field action reporting** – field officers will be able to upload action(s) taken at the scene, details of which will be shared by the initiator to ensure transparency in reporting. The officer will be guided to the scene by the reporter's location.
vi. **Notification** – the system should notify users when there is an action that requires their intervention.
vii. **Feedback** – the member of the public who made the report will be able to give feedback on the level of satisfaction with an attending officer's service delivery for every incident reported.
viii. **Appraisals** – an officer's supervisor and peers will be able to rate the officer where applicable.
ix. **Score generation** – the system should be able to aggregate all scores relating to an individual officer. Authorised users will be able generate and view individual officers' appraisal scores computed by the algorithm. The Fuzzy Inference System should be able to classify an officer's performance based on the predefined rules.
x. **Reports** – the system should be able to organise data within the database and display records and reports when required.

4.2.2 Non – Functional Requirements

Non-functional requirements include system properties that will determine the level of satisfaction by various users. Some non –functional requirements that were considered include:

i. **Performance** – the system must be able to quickly process requests to enable speedy reporting of incidents as well quick feedback. Slow response may make the users shun away from giving feedback. It also should allow multiple users without degrading performance.

ii. **Scalability** – the system should be easily modified whenever additional functionality is desired. Errors should also be identified and resolved with ease.

iii. **Accuracy** – the ratings should be calculated with precision to generate fair scoring for every officer.

iv. **User friendliness** - the interface must allow easy navigation by all users.

v. **Security** - the system should be secure to protect against unauthorised access. It should also protect the integrity of any data uploaded.

vi. **Reliability** - the system should always be available and should not have unplanned downtime.

4.3 System Architecture

The architectural design of the appraisal model is as shown on Figure 5.1. The system will adopt a 3-tier architecture with a Presentation Layer, Application Layer and Database. The presentation layer will comprise the mobile and web application and will provide users the interface for data input and records viewing. The Application layer will perform the score calculations from data inputs, run the fuzzy logic controller and perform data storage functions continuously as long as there is fresh input data from the presentation layer. The database receives and stores information obtained from inputs and processes and also permits access of reports to authorised users. The design is as shown on Figure 4.1.

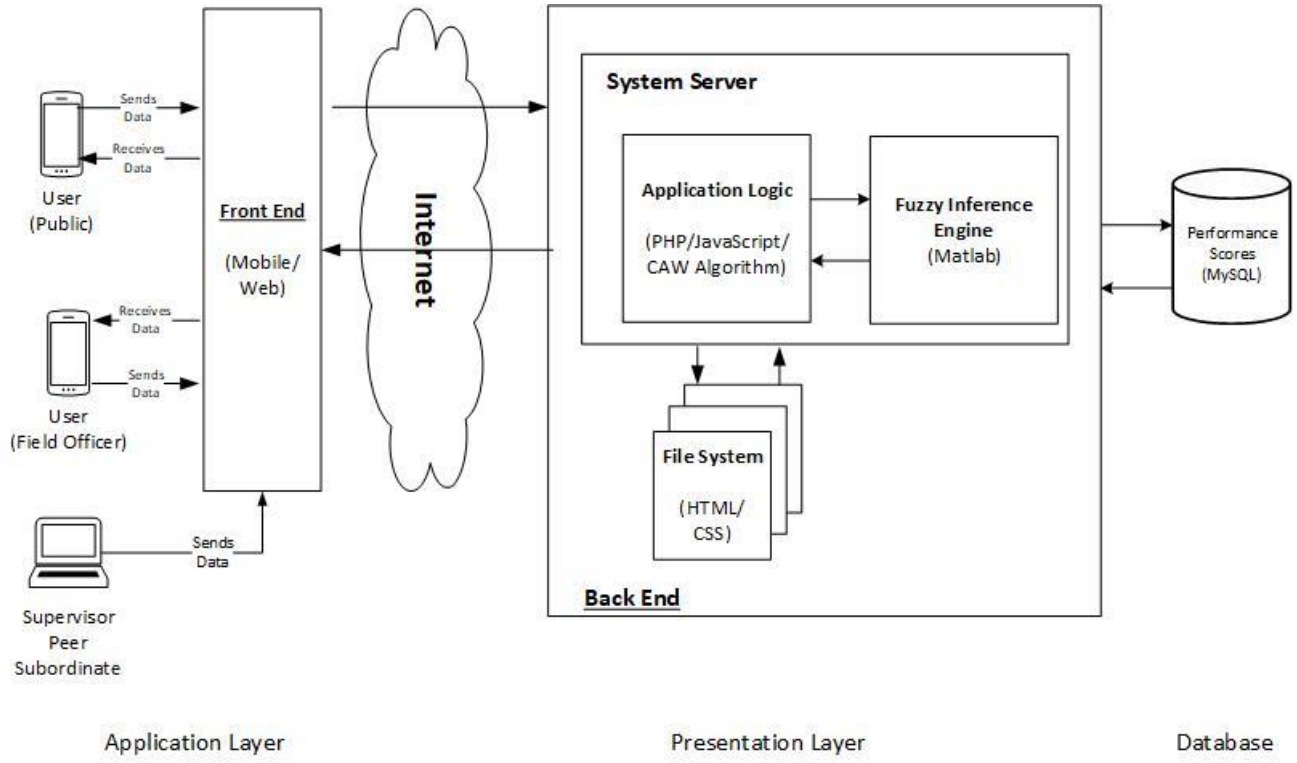

Figure 4-1 System Architecture Diagram

4.4 System Modelling

The system and user requirements were modelled using diagrams that illustrated the relationship between actors and the use cases.

4.4.1 Use Case Description

The main use cases representing the system's events when actors interact with system are described below:

Table 4.1 Use Case - File Report

Use Case	File Incidence Report
Description	The user makes a report indicating predefined category
Primary Actor	User

Pre-condition	Must have logged into the system
Post condition	Connectivity is established.

Actor	**System**
1. User accesses mobile interface	
	2. System displays interface with category dropdown and details field
3. User selects category, gives brief description and sends report.	
	4. System saves report and assigns officer.

Table 4.2 Use Case - Give Feedback

Use Case	Give Feedback
Description	The user gives feedback regarding service offered by officer.
Primary Actor	User
Pre-condition	Must have filed a report via the mobile client.
Post condition	Report details availed to user.

Actor	**System**
1. User accesses new prompts.	
	2. System displays feedback request form.
3. User completes feedback request form.	
	4. System saves report and computes score.

Table 4.3 Use Case - File Field Report

Use Case	File field report
Description	The officer uploads findings from information gathered on the scene.
Primary Actor	Officer
Pre-condition	Must be registered in the system and have a case assigned to them.
Post condition	Backend connectivity established.

Actor	**System**
1. User accesses assigned complaint and uploads report.	
	2. System saves report and prompts complainant to give feedback.
3. User selects category, gives brief description and sends report.	

Table 4.4 Use Case - Upload annual appraisal

Use Case	Upload annual appraisal
Description	The supervisor makes annual appraisal using prescribed format.
Primary Actor	Supervising Officer
Pre-condition	Must be logged in registered user.
Post condition	Subordinate officers' database accessed.

Actor	**System**
1. User accesses web based backend and selects officer.	
	2. System displays standard appraisal form.
3. User fills in active field and submits report.	
	4. System saves report and aggregates with feedback score.

4.4.2 Use Case Diagram

The use case is a representation of the different points of contact with the system. It is a relationship between players (actors), the methods that the system will use to access data and background operations. The actors include the customer (Member of the public), the attending officer, the supervising officer, the system administrator and the system. This is as shown in Figure 4.2.

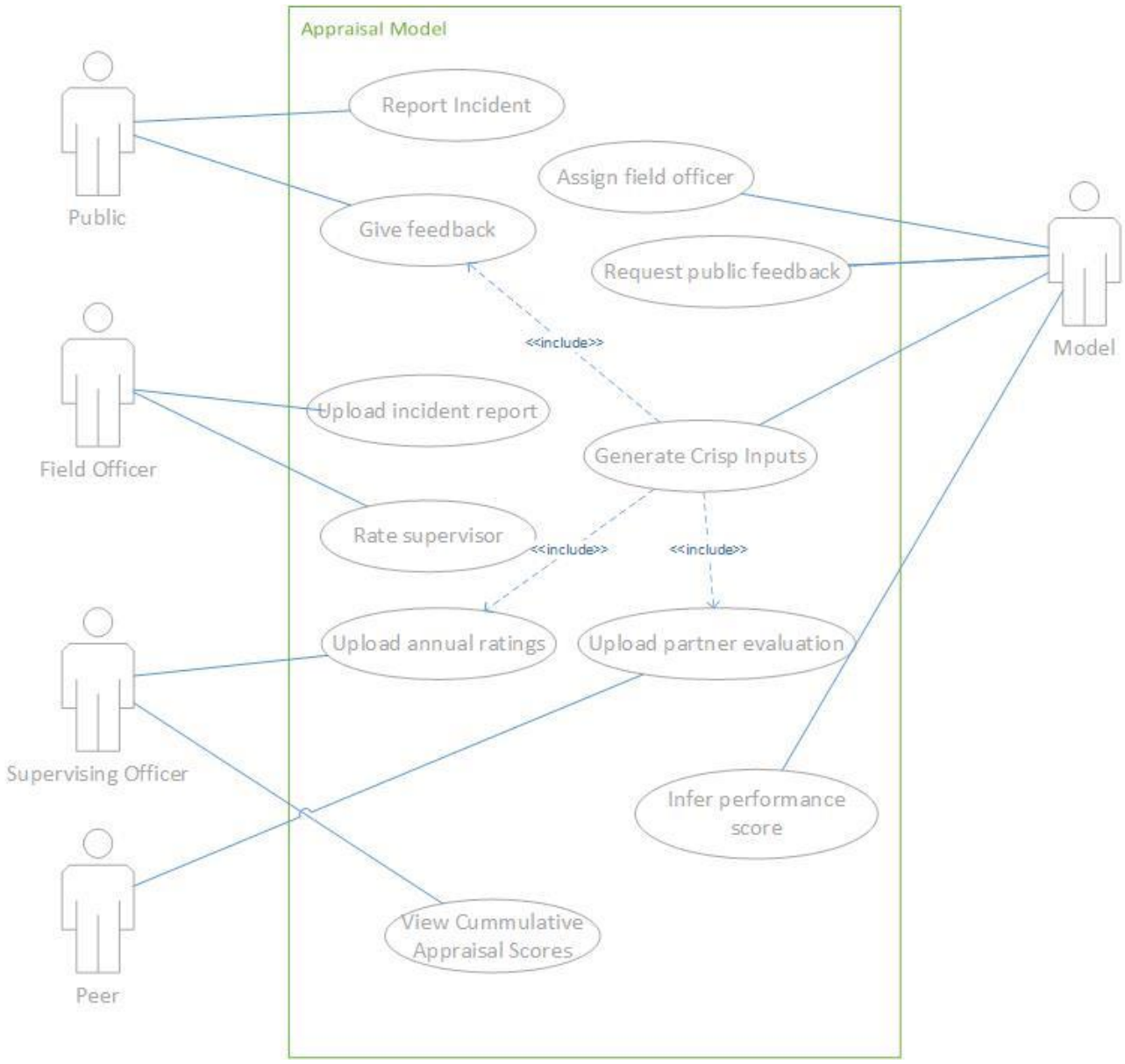

Figure 4-2 Use Case Diagram

4.4.3 Sequence Diagram

The sequence diagram depicts the sequential flow of events between various actors. The member of the public begins the sequence of activities by reporting an incident requiring the attention of a Police Officer. Based on the nature of the incident reported, the system prompts an appropriate officer who processes the complaint and reverts to the system. The system then requests the customer to rate the service provided after which the scores are stored. These form crisp inputs to the Fizzy Inference Controller that uses a rule base to arrive at a performance score. The sequence diagram in Figure 4.3 represents the interaction between the user and the various functions that are centric to the running of the system and generation of the performance score.

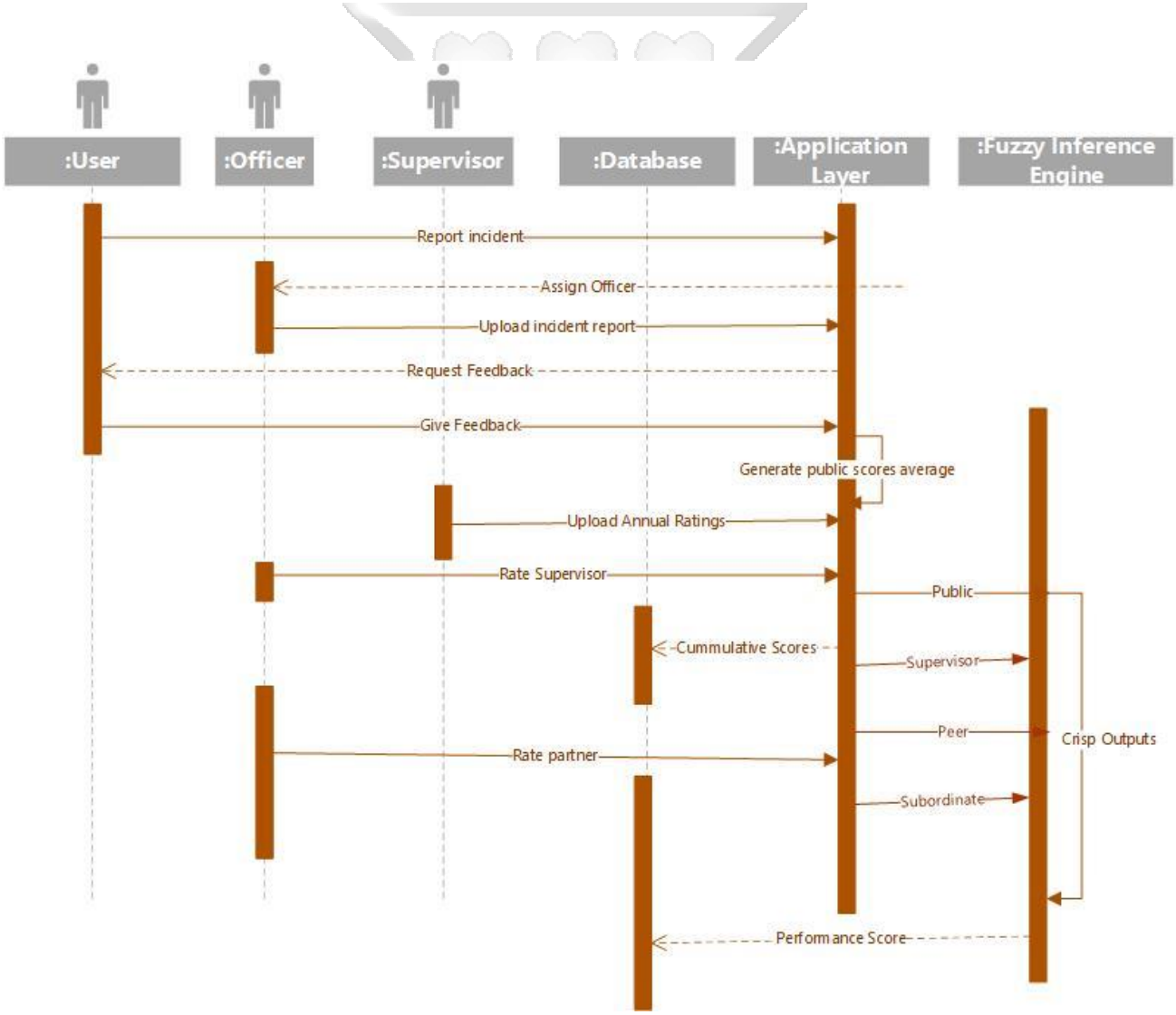

Figure 4-3 Sequence Diagram

4.4.4 Context Level Diagram

The context level diagram defines the interaction between the system and the main actors and defines the system boundaries. The main system actors are the customer, attending officer, supervising officer and the system administrator.

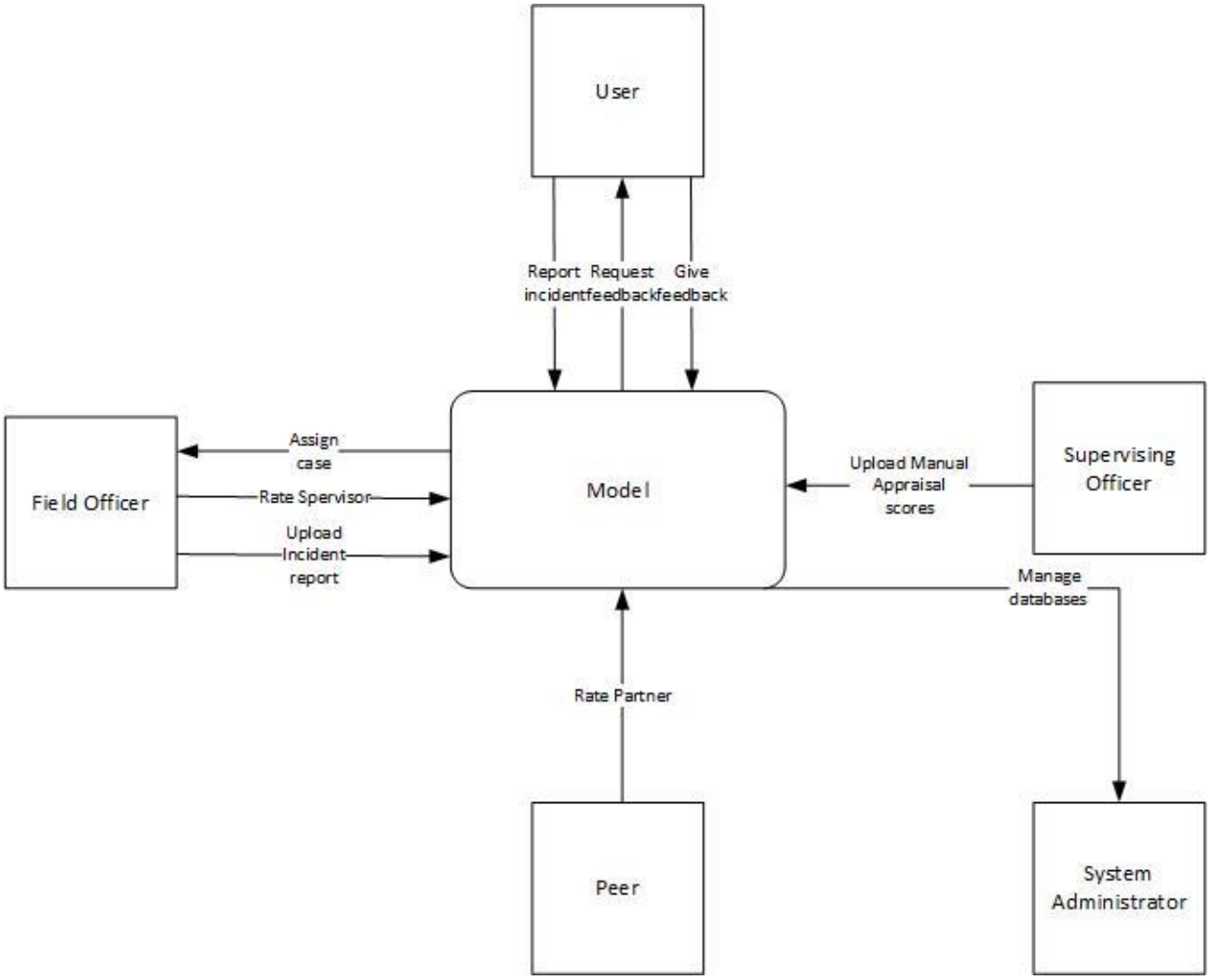

Figure 4-4 Context Level Diagram

4.4.5 Level 1 Data Flow Diagram

The level 1 DFD illustrates a breakdown of the context level diagram giving further detail of information flow between the system actors. It also shows where data resulting from various system processes will be stored.

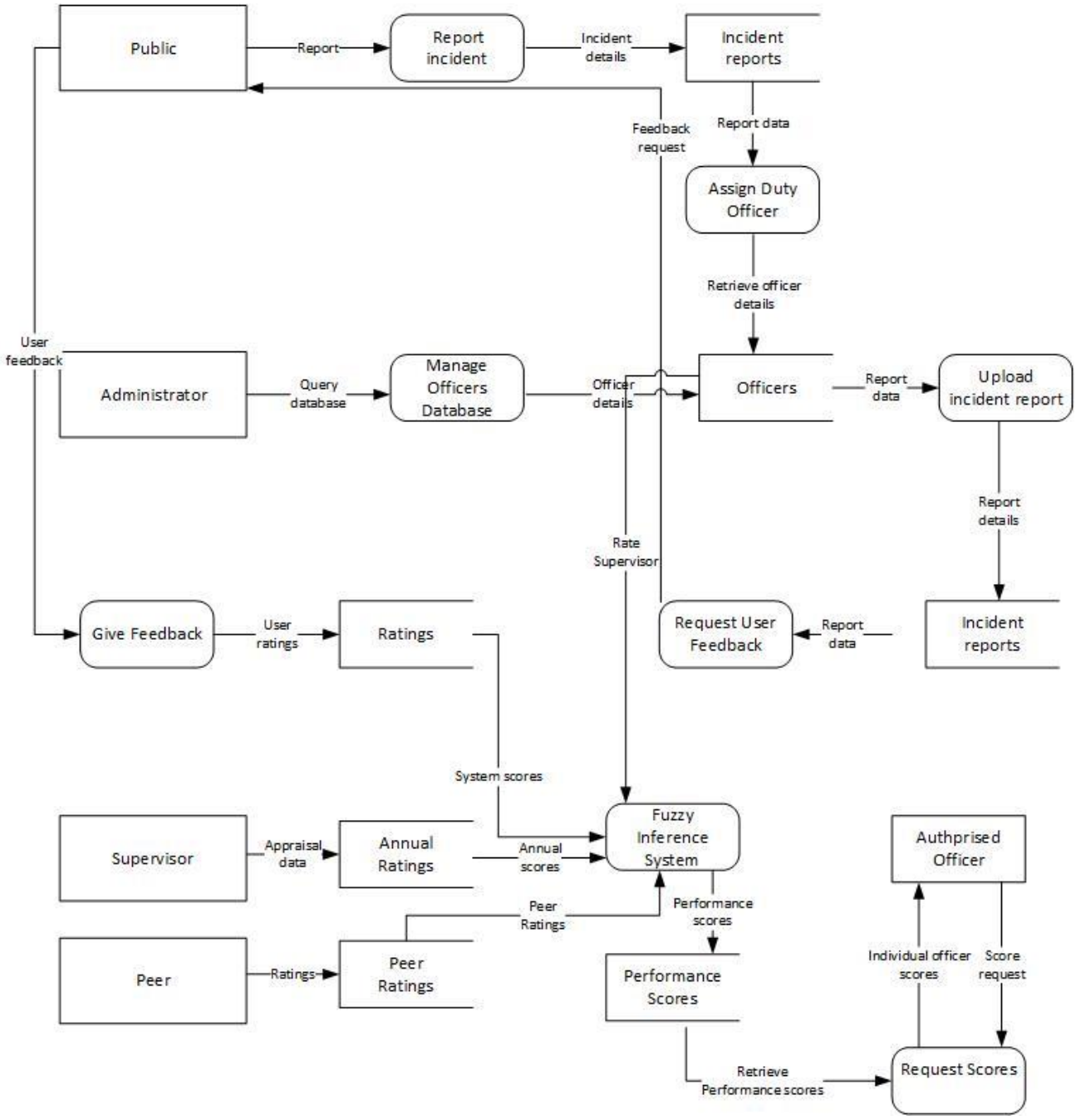

Figure 4-5 Level 1 Data Flow Diagram

4.4.6 Entity Relation Diagram

The entity relation diagram demonstrates the logical model of the database. This will be helpful in understanding the flow of data between different objects in the system as well as how best to model the data to fit within a specific use case. The database entity-relations are as presented in Figure 4.6.

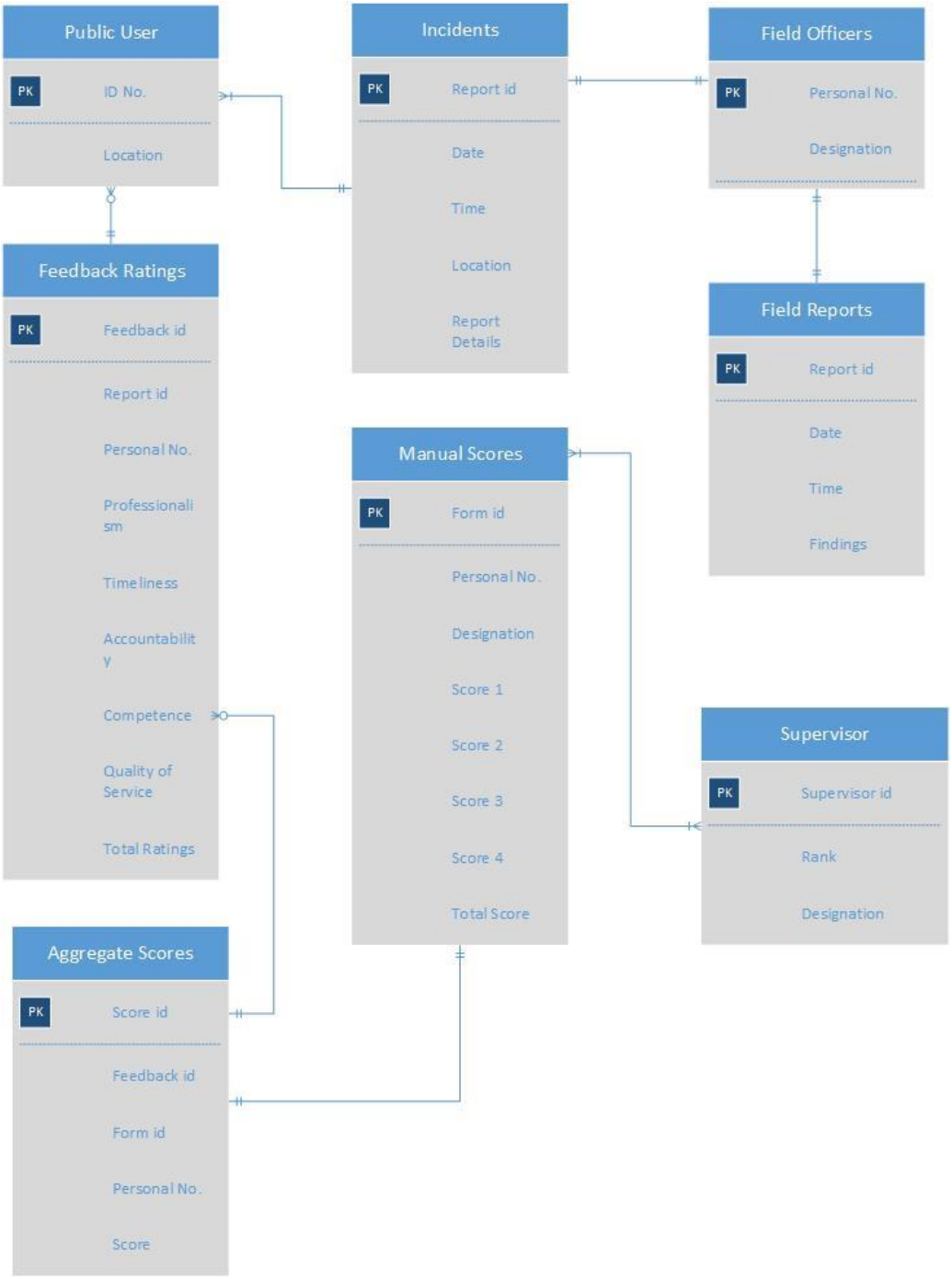

Figure 4-6 Entity Relation Diagram

Chapter 5: System Implementation and Testing

5.1 Introduction

This chapter describes the implementation of the system design to a working prototype. The system will be composed of three main parts; an android based client side that will provide the reporting and feedback functionality, a web based server side from which system administration, manual scoring and reporting will take place and a relational database. Prototype testing and user acceptance testing will be discussed at the end of the chapter.

5.2 System Components

The system output will be appraisal scores generated from an aggregate of rating from customer ratings and an officer's immediate supervisor through an algorithm that will ensure consistency and fairness in rating. The system will adopt a 3-tier architecture. The presentation layer (mobile-based and a web based client side), an application layer through which score calculations will be achieved through the CAW algorithm and a database. The hardware requirements include a mobile device running Android version 6 and above, a personal computer running Windows 7 or later, Relational Database Management System (MySQL 5.0.45), HTML, JavaScript, Web Application Server (Apache) and an Internet Browser (Chrome or Firefox)

5.2.1 The Mobile Client Side

The mobile-based front end will allow the users (customer and officer) to interact with the system to provide input data.

1. System sequence will be initiated when an already registered user logs into the app and reports an incident, which will be stored in the incident's database on the server side.
2. The system selects from the officer's database an appropriately deployed officer based on the category of incident selected by the user (e.g. traffic).
3. The officer will attend to the complaint and log the findings back to the system through their mobile interface, which will have this additional functionality from the customer's view.
4. Upon receipt of the officer's findings, the system will prompt the concerned customer to rate the officer through a set of questions meant to evaluate an officer's professionalism, accountability, timeliness, competence and quality of service.

5. The officer's efficiency will be measured by the turn-around time between assignment of the responsibility by the system, to arrival at the scene and uploading the findings back to the system.
6. The officers' view will also allow them to view ratings from every single source as well as the aggregated appraisal score.

The proposed evaluation dataset that will ensure uniformity in rating will be as described in the table below:

Table 5.1 Rating Dataset - Public

	Variable	Question	Possible values
1.	Timeliness	Did the officer respond in a timely manner?	Exceptional (5), Neutral (3), Very Poor (1)
2.	Professionalism	Was the attitude and demeanour of the attending officer a display of professionalism?	Exceptional (5), Neutral (3), Very Poor (1)
3.	Accountability	Did the officer give useful feedback regularly?	Exceptional (5), Neutral (3), Very Poor (1)
4.	Professionalism	Did the officer go out of their way to help resolve the complaint?	Exceptional (5), Neutral (3), Very Poor (1)
5.	Competence	To what extent have the officer's services inspired your confidence in the competence of the National Police Service?	Exceptional (5), Neutral (3), Very Poor (1)
6.	Quality of Service	To what extent are you satisfied with the manner in which the officer handled the case?	Exceptional (5), Neutral (3), Very Poor (1)

5.2.2 The Web-based Client Side

Data from incident reporting, officer field report and customer ratings were uploaded to the localhost server. The backend interface was built using HTML, PHP and MySQL served as the storage database. This is the core-subsystem that is responsible for generating appraisal scores, managing users and presenting the ratings to authorised users. This interface enables database management by the administrator. The supervising officer and a selected peer gets rights to login to the backend where

they can access an officers profile in which there is a the standard public officer annual performance appraisal form with agreed performance targets, expected performance indicators, mid-year review remarks and performance scores as shown below:

Table 5.2 Rating Dataset - Supervisor

	Agreed Performance Targets/Specific Tasks Assignment	**Expected Performance Indicators**	**Mid-Year Review (remarks)**	**Performance Appraisal Score (1 to 10 with 1 being very poor, 5 average and 10 Excellent)**
1.	Crime Resolution	No. of crime reports		
2.				
3.				
4.				
5.				
Total appraisal score on performance targets (Max. 50)				
Mean appraisal score (%)				

The supervising officer will also be able to view the historical performance for officers placed under their charge. The same will apply to authority mandated to manage the human capital within the Police Service. The system administrator will be able to manage users, administer the databases and access error logs to enable timely maintenance interventions.

The peers will be asked to give ratings as follows:

Table 5.3 Rating Dataset - Peers

	Please rate your partner along the following lines:		
	Variable	**Question**	**Possible values**
1.	Timeliness	Does your partner regularly report on time shifts and wait to be relieved before leaving the duty station?	Exceptional (5), Neutral (3), Very Poor (1)
2.	Professionalism	Does your partner act professionally in dealings with clients and colleagues?	Exceptional (5), Neutral (3), Very Poor (1)
3.	Accountability	Does your colleague learn from their mistakes and take responsibility to correct them?	Exceptional (5), Neutral (3), Very Poor (1)
4.	Competence	Does your partner apply their training and on job learning to address the unique client needs related to policing?	Exceptional (5), Neutral (3), Very Poor (1)
5.	Quality of Service	Is your partner focussed on the needs of the client?	Exceptional (5), Neutral (3), Very Poor (1)

Where applicable, officers could be asked to rate their supervisors as follows:

Table 5.4 Rating Dataset – Subordinate

	Place rate your supervisor along the following lines:		
	Variable	**Question**	**Possible values**
1.	Timeliness	How would you rate response to facilitation requests?	Exceptional (5), Neutral (3), Very Poor (1)
2.	Professionalism	Are conflicts handled in an appropriate manner?	Exceptional (5), Neutral (3), Very Poor (1)
3.	Accountability	Performance feedback is given regularly?	Exceptional (5), Neutral (3), Very Poor (1)

4.	Competence	Does their work supervision mechanism portray competence in leadership?	Exceptional (5), Neutral (3), Very Poor (1)
5.	Quality of Service	To what extent are you satisfied with the supervision level?	Exceptional (5), Neutral (3), Very Poor (1)

5.2.3 Data Processing

The cumulated average weighting (CAW) algorithm was used to aggregate appraisal scores from the two sources. The naïve-cumulated average-weighting algorithm can be used to get the variance in a set of appraisal scores for an individual (Ahsan, 2016). The simplistic pseudo code for this algorithm is as shown in Equation 5.1.

- **Let ← 0, SumSq ← 0**
- **For each datum x:**
  - **n ← n + 1**
  - **Sum ← Sum + x**
  - **SumSq ← SumSq + x * x**
- **Var = (SumSq – (Sum * Sum)/n)/(n – 1)**

Equation 5.1: Pseudocode for Naïve CAW Algorithm

The equation behind the Naïve CAW is as shown in Eqn. 5.2

$$\sigma^2 = (\overline{x^2}) - \overline{x}^2 = \frac{\sum_{i=0}^{N} x_i^2 - (\sum_{i=1}^{N} x_i)2 \ /N}{N}$$

Equation 5.2: The Naïve CAW Algorithm

The Naïve CAW is further advanced into the two-pass cumulative average-weighting algorithm. This is the algorithm that was applied to aggregate an officer's scores reported by a member of the public. It works in the following series of steps;

a. The algorithm first computes the sample means of the appraisal results from the responses given by the public then computes the squares of the difference of the means to get the standard deviation. From the standard deviation, the average appraisal score

can be determined from a set of unarranged data elements. This is as shown in Equation 5.2

$$\overline{x} = \frac{\sum_{j=1}^{n} x_j}{n}$$

Equation 5.3: Computation of sample means

b. The algorithm then follows a series of loops to tally and include results of appraisal from the public incident reports and the peer results offered by the other police officers appraising the individual officer. The pseudocode for the process is as shown in Equation 5.4.

```
def two_pass_variance(data):
        n = sum1 = sum2 = 0
        for x in data:
                n + = 1
                sum1 += x
        mean = sum1/n
        for x in data:
                sum2 += (x – mean) * (x – mean)
        variance = sum2/ (n - 1)
return variance
```

Equation 5.4 Pseudocode for implementing the two-pass CAW algorithm

In both algorithms, the variable (x) represents the inputs that are required for the system. The value (n) represents the number of times the algorithm will iterate summing up new appraisal scores for the officers, with each incident responded to by the residents. The system takes inputs (x) from different sources; the public and the police supervisors. After the public appraises the officer, the value (x) is assigned to a particular pool of data, represented by an array. The same officer then gets a supervisor rating, which further increases the value of x. It is these scores that can then be viewed by an authorised officer for purposes of human capital management decision making.

c. From the Appraisal variance, computed by the appraisal scores of the public feedback, the supervisor, peer and where applicable the sub-ordinate, the algorithm combines additional x inputs from both sources to get the final aggregated appraisal score.

5.2.4 The Database

The relational database was implemented using MySQL to facilitate creation, extraction, analysis and storage of provisional ratings from the public, supervisors and peers as well as the aggregated appraisal scores. Data is stored in plain text while passwords are encrypted using MD5 algorithm. The mobile and web application all connect to a single database which will be backed up regularly.

5.3 System Development

In developing the system, emphasis was placed on having accurate data input from the client side as well as real time update of the databases on the server side.

5.3.1 Mobile Client Side

The mobile front end was developed on Android Studio Version 3.3.1 on Java code and the query language for connectivity with web based back-end implemented through JSON. The user interface and logic was developed using XML. The primary users of this front end will be the Public and Field Officers. Both will have the same interface but the Field Officers will have extra functionality to enable filing of field reports. Selection of an incident category will inform the officer selected to attend to the scene.

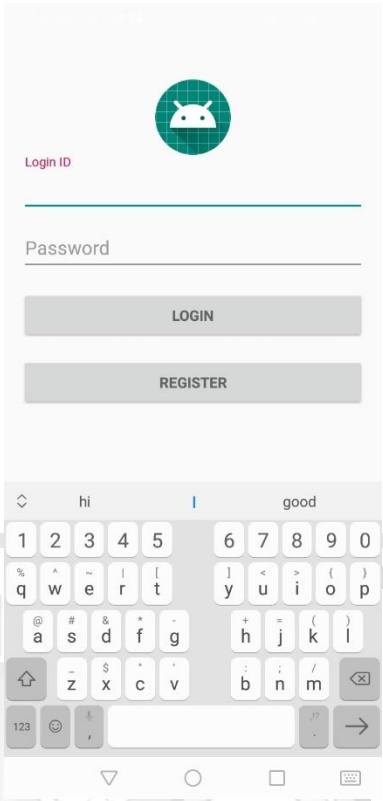

Figure 5-1 Mobile Front-end Home Screen

5.3.2 Web Application

The web application was developed on PHP using the CodeIgniter framework. The views are generated through JavaScript, HTML5, CSS3 and JQuery. The system administrator will have access to the entire backend and can access the Registered Users, Reports, Ratings and the Appraisal Scores:

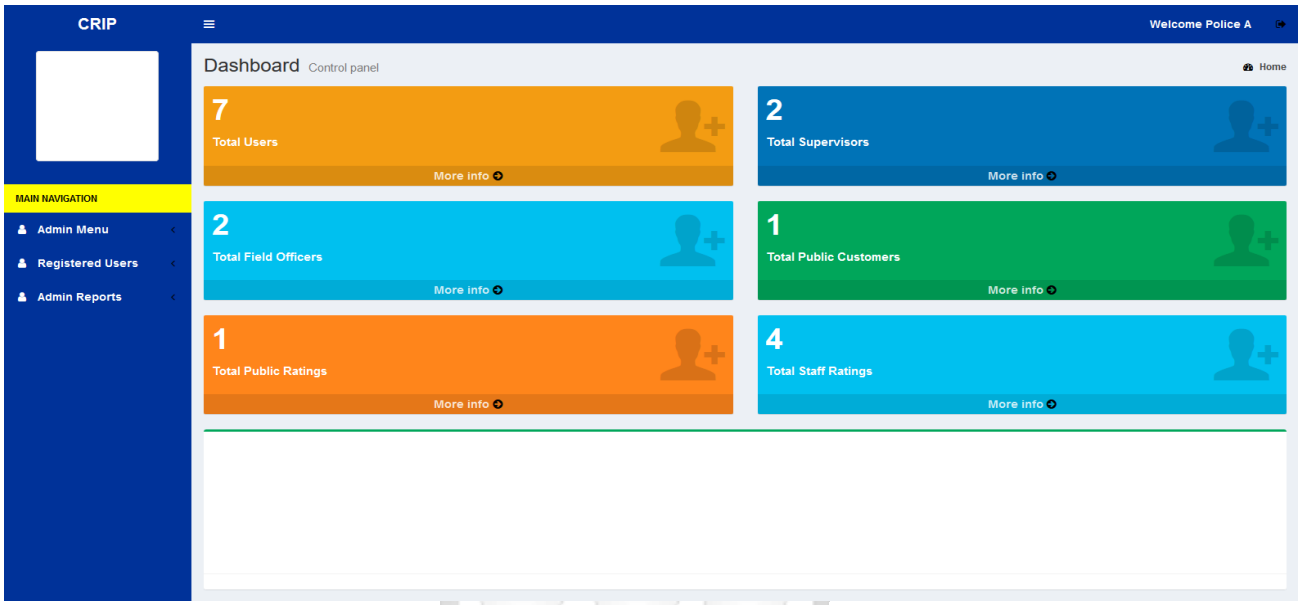

Figure 5-2 Server Side Administrator Home Screen

The public view will include only the ID and mobile phone numbers for the sake of the users remaining anonymous while the officers' views (Field Officer, Peers and Supervisors) will include their Personal Numbers, Names, Rank and Designation:

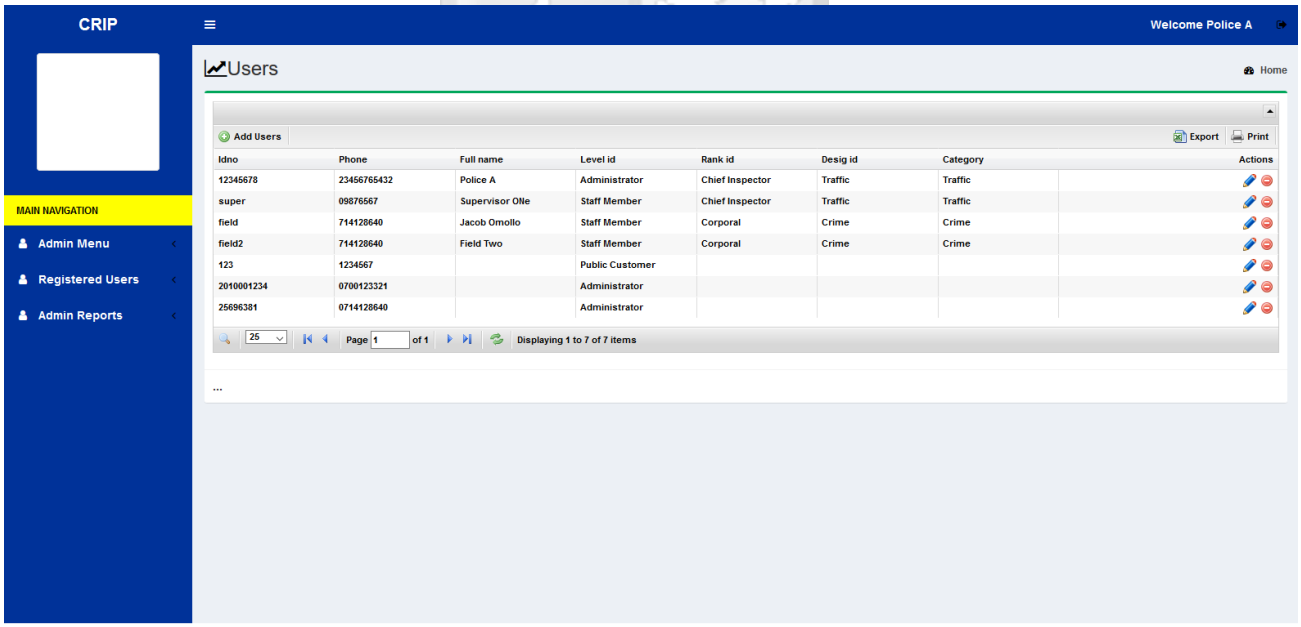

Figure 5-3 Server Side Users

The officers' view will include a provision for the administrator to add or remove users:

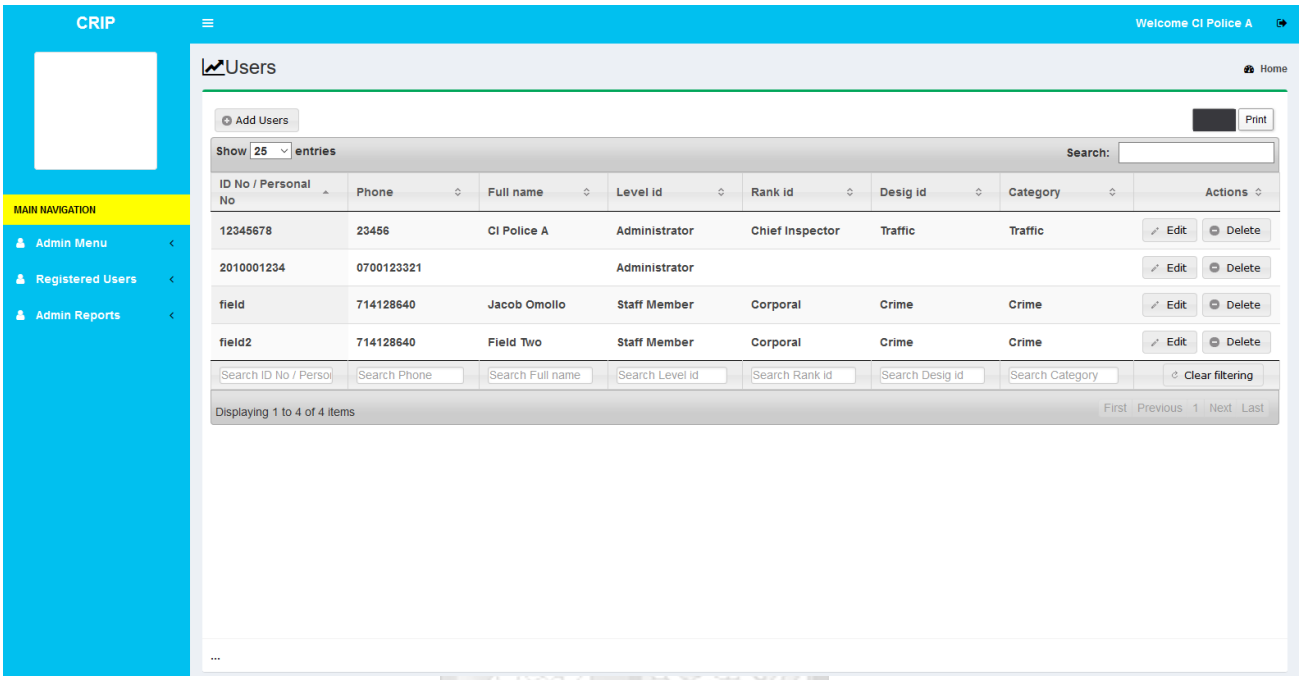

Figure 5-4 Server Side Field Officers View

The reports view will have tables detailing feedback from the public and field officers:

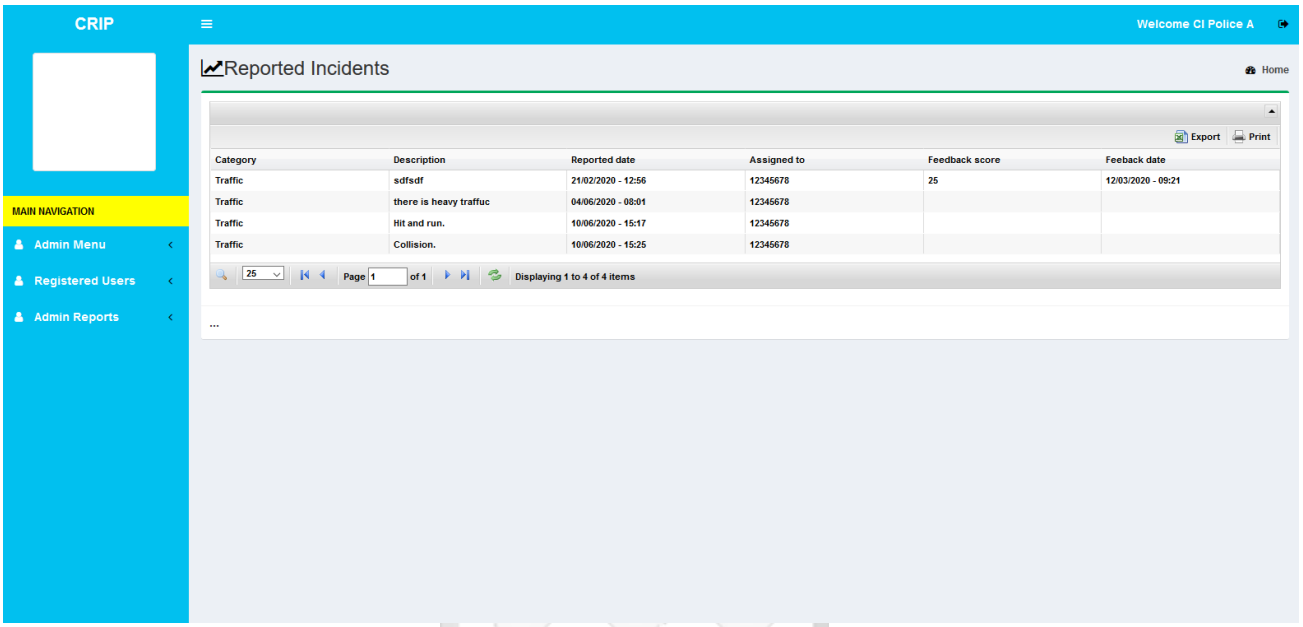

Figure 5-5 Public Reports View

Ratings from all the raters can be viewed individually as the following example for the public:

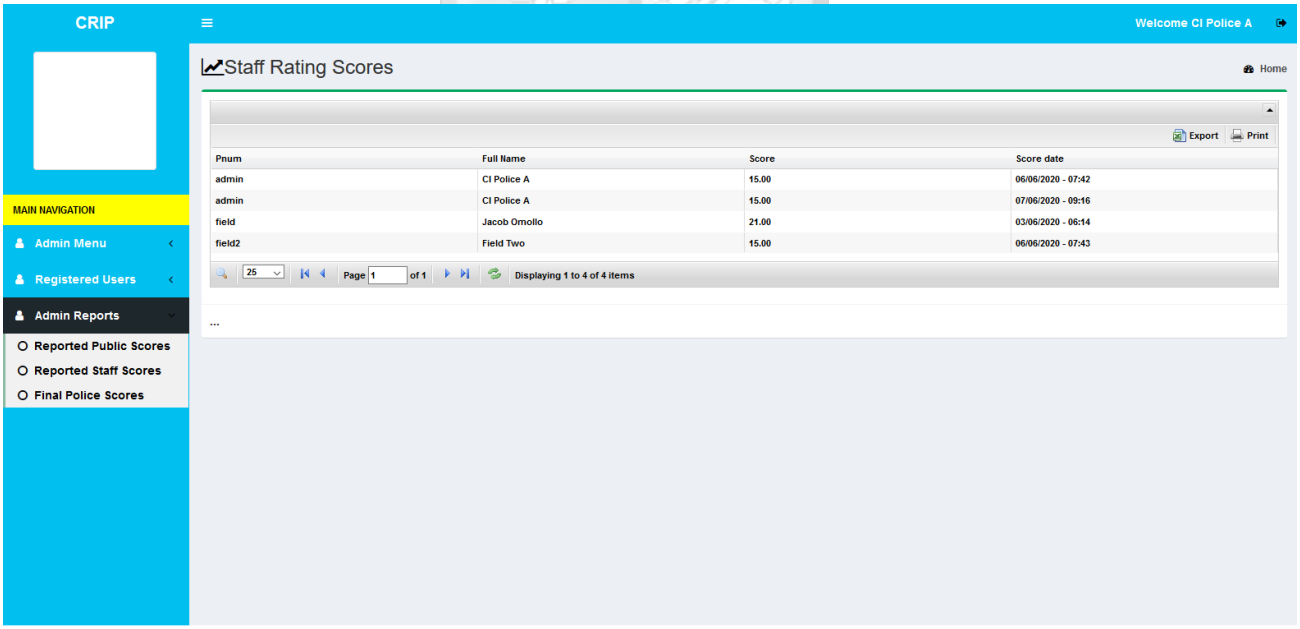

Figure 5-6 Public ratings view

The final score arrived but by the algorithm computation can be viewed by authorised users as follows:

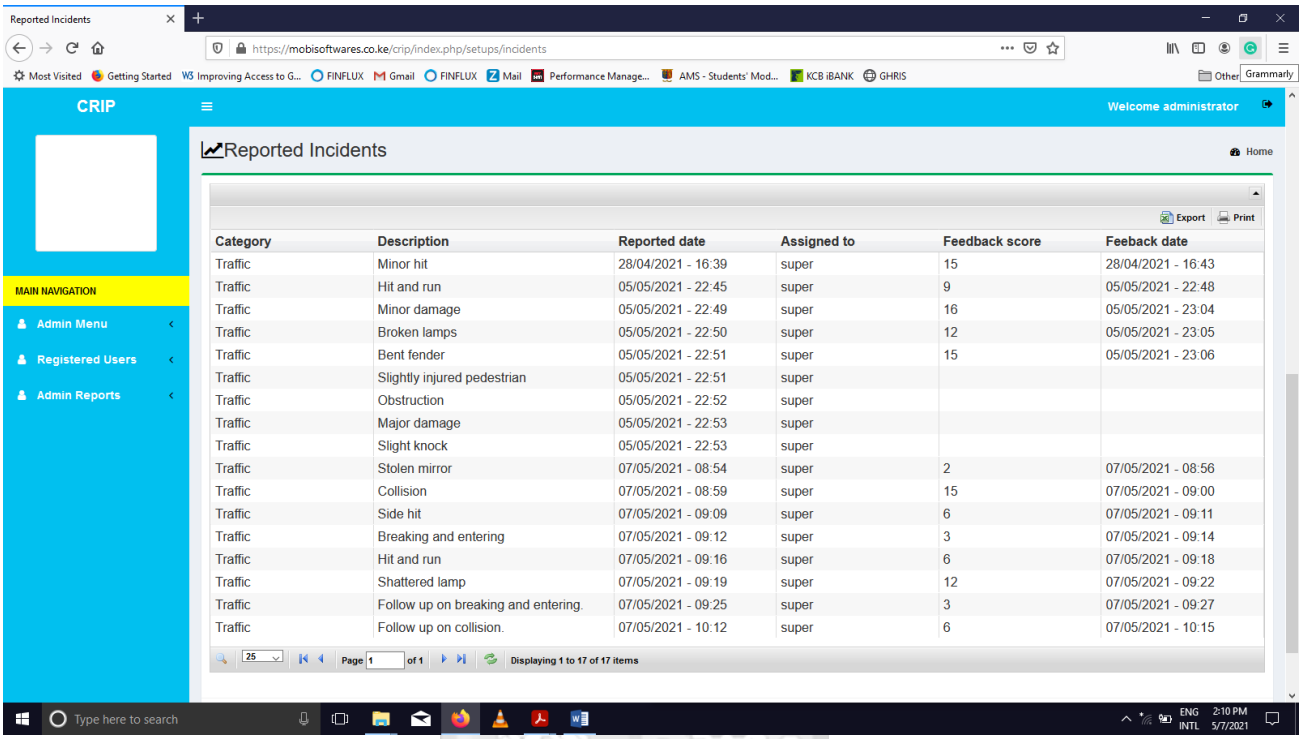

Figure 5-7 Public Appraisal Scores

Test data was fed into the system and completed a cycle from the moment a member of the public makes a report, a field officer uploads a report from the scene and the person reporting giving a feedback on the field officer was conducted (Fig. 5.9). To determine appraisal rating, testing on public feedback was conducted to determine a range of values within which an officer would be rated as weak, average or good. A flat rating of three (3) would be suggestive of a neutral attitude from the respondent resulting in an overall invalid rating that will be disregarded. The rating scales that will apply for each variable in Table 5.1 are as follows:

Table 5: Appraisal Rating Score Scale

	Scores Range	**Appraisal Rating**
1.	$2 \geq$ Score < 6	Weak
2.	$6 \geq$ Score < 12	Average
3.	Score $= 3$	Invalid

4.	12 ≥ Score ≥ 15	Good

5.4 Fuzzy Control System

Fuzzy logic handles uncertainty much better than mathematical formulas. A fuzzy controller can approximate the relationship between inputs and outputs via interpolation in a vague environment. This research employed fuzzy logic to minimise subjectivity in the appraisal process. Fuzzy logic can handle problems with imprecise or incomplete data, giving a better degree of accuracy as compared to binary which gives a more realistic model of the process. There are individuals who would score Very Good (1) or Very Poor (0) but fuzzy logic allows us to define the whole range of values in between.

An appraisal system built with a rule based fuzzy controller will also make it easy for an administrator to determine performance in real time as long as a set of inputs is updated. For example, as an officer is rated by members of the public, the administrator can give more prominence to this input in the rules base and arrive at a decision with speed and accuracy.

The fuzzy set in this research will be the Performance Score while the fuzzy subsets will be {Very Poor, Poor, Average, Good, Very Good}.

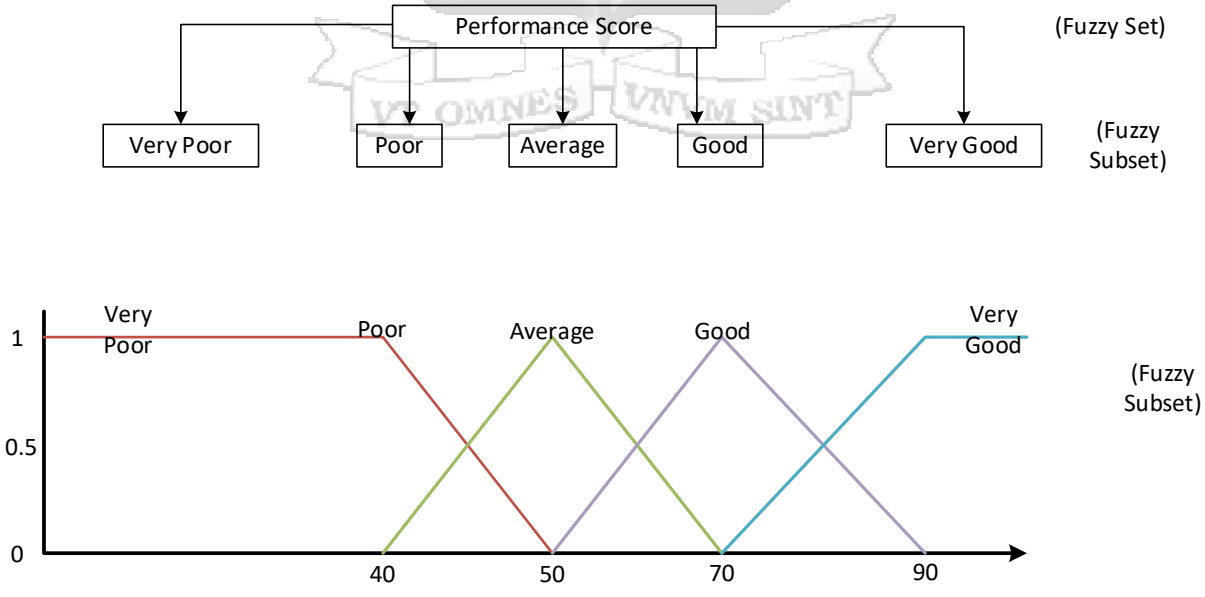

Figure 5-8 Fuzzy Set and Sub-sets

5.4.1 Fuzzy Logic Description

Fuzzy logic defines the degree to which a statement is true and resembles human reasoning. An element in a fuzzy subset can belong to more than one set because of overlapping membership functions of sets. The general structure of the proposed fuzzy appraisal model is as follows:

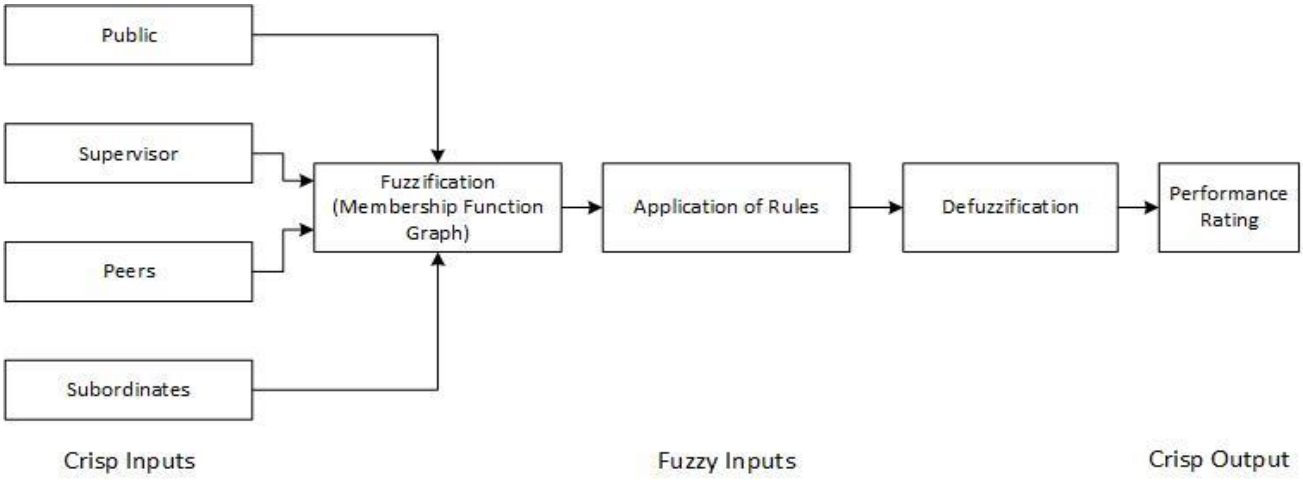

Figure 5-9 Proposed FIS for Performance Appraisal Module

5.4.2 Fuzzy Logic Toolbox

The Fuzzy Logic Toolbox was chosen to develop the model in this study. According to Mathworks, this toolbox is ideal for the following reasons:

i. It supplies a fuzzy inference engine that can execute a fuzzy system as a stand-alone application.
ii. It provides an easy to use interface for applying modern fuzzy logic techniques.
iii. It allows the user to perform fuzzy logic control and identification.
iv. It provides the ability to use fuzzy logic when appropriate with other control techniques.
v. It provides the ability to generate code for various uses.

5.4.3 Structure of the Fuzzy Logic Model

The main steps in the model are as follows:

Step 1: Fuzzification of inputs

Fuzzification determines membership values for the given crisp inputs. The fuzzified value will range between 0 and 1. Membership was determined by intuition approximating peak points for the fuzzy subset since designing the fuzzy controller requires knowledge on the real world behaviour of the expected output. Membership functions make it possible to quantify a fuzzy variable and represent it in a fuzzy set graphically. The fuzzy membership graph for each crisp input given triangular shaped membership functions will be as below:

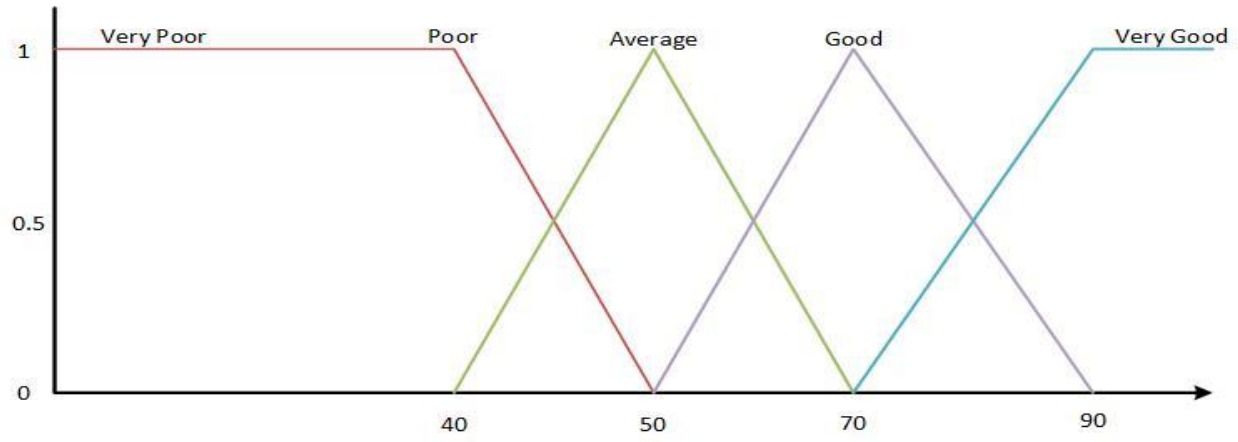

Figure 5-10Fuzzy Membership Graph

Step 2: Rules Evaluation

Fuzzy logic does not work without rules. Setting fuzzy rules a very important step in the design and implementation of a fuzzy control system. The Fuzzy Inference System (FIS) carries the model's intelligence. The fuzzy operator is applied to obtain one value given that every crisp input will have two memberships. To ease the rule development process, a Fuzzy Associative Matrix (FAM) as sampled below:

Table 6 Fuzzy Associative Matrix

Rule	**Inputs**					**Output (Performance Score)**
1	Very Good	Good	Average	Poor	Very Poor	Good
2	Good	Very Good	Good	Average	Poor	Good
3	Average	Poor	Very Good	Good	Average	Average
4	Poor	Very Poor	Average	Very Good	Good	Average
5	Very Poor	Good	Very Good	Good	Average	Average

The more the set values the more complex the rule base. Matlab was used in this research to develop the fuzzy rule base. Some rules may be theoretically possible, but practically they do not make sense. Some will be applicable to a given scenario meaning the particular rule will fire and some will not fire.

Fuzzy logic has no capacity for learning therefore the rules can be developed from experts (intuition), from data (clustering, genetic algorithm) or using software like Matlab to generate evolving fuzzy rules.

The calculated score forms the inputs (antecedents) and the appraisal rating the output (consequents). Fuzzy rules will then be used to map the input space to an appropriate output space.

The fuzzy model with multiple inputs was implemented as follows:

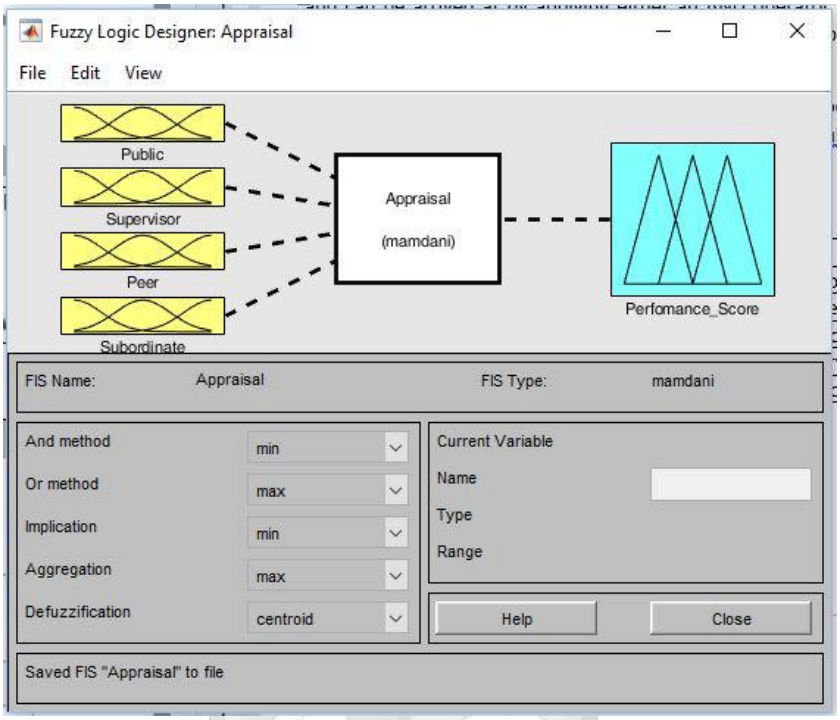

Figure 5-11 Fuzzy Logic Design (Matlab)

The triangular membership functions were constructed then the rule base for possible values developed as sampled below:

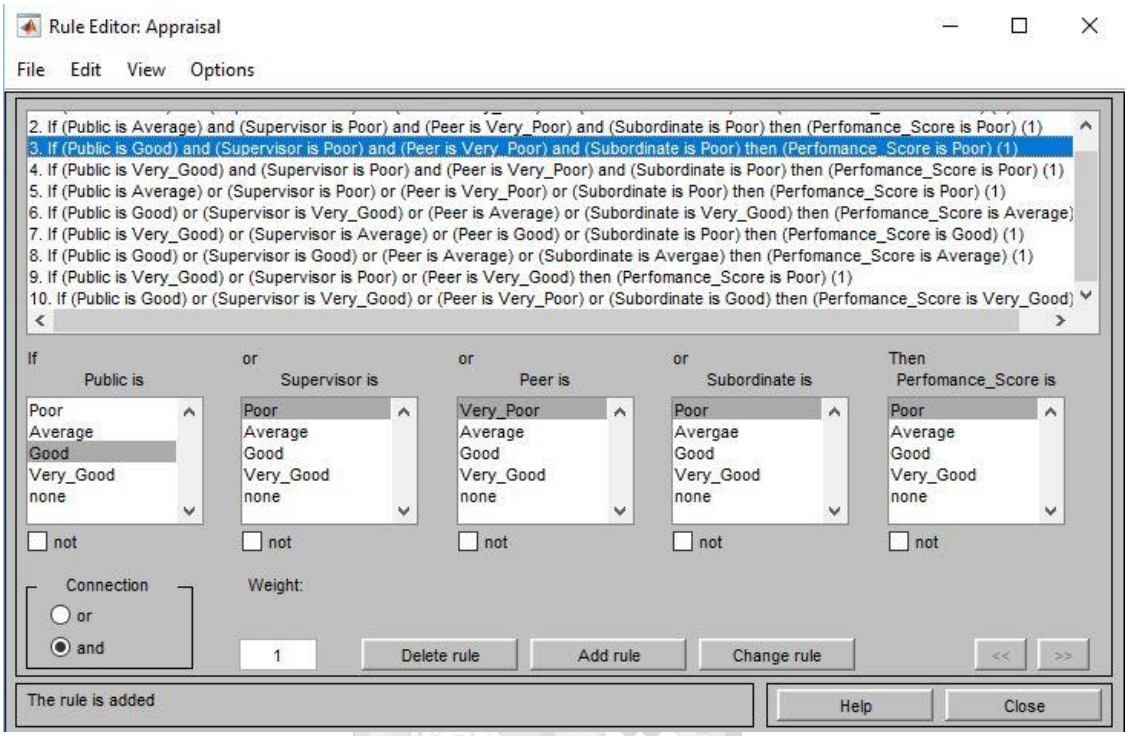

Figure 5-12 Sample Matlab Fuzzy Rule Generator

Step 3: Defuzzification

The last step is to convert the fuzzy outputs into crisp values. The most common and accurate method is the centroid technique. The linguistic result will be displayed as per the rule invoked.

To illustrate the cycle from inputs to outputs by experiment, scores were selected as 43 and 65 for two inputs. These scores will have a degree of membership to either the average or good category illustrated as follows:

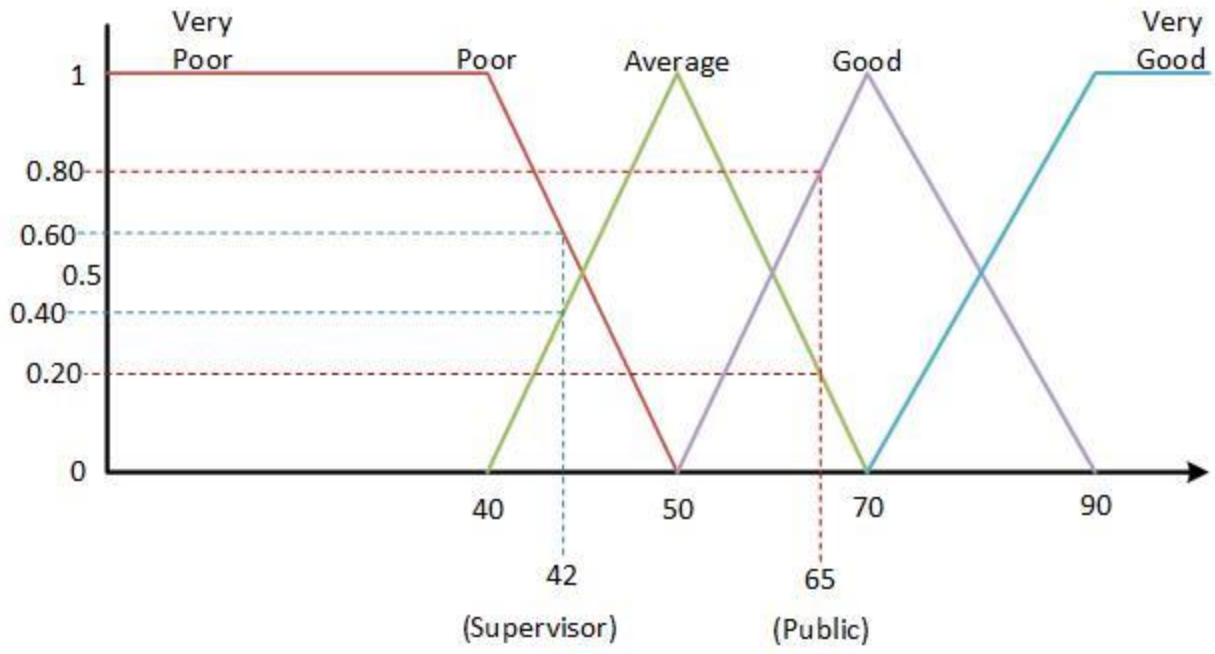

Figure 5-13 Fuzzy Outputs (Public and Supervisor)

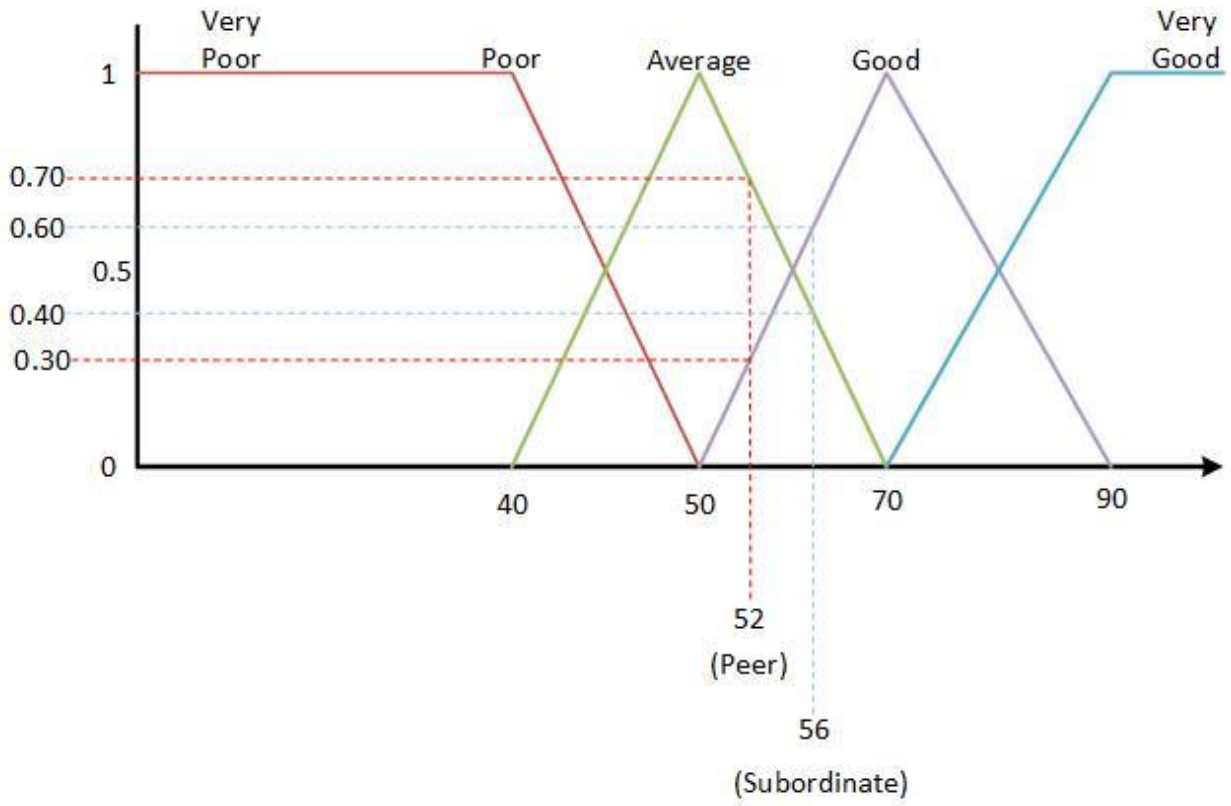

Figure 5-14 Fuzzy Outputs (Peer and Subordinate)

The dotted lines map the crisp scores to the possible fuzzy outputs. The fuzzy outputs are tabulated as follows:

Table 7 Fuzzy Outputs

	Fuzzy Subset					
Input	**Very Poor**	**Poor**	**Average**	**Good**	**Very Good**	**Output (OR Operator)**
Public	0.00	0.00	0.20	0.80	0.00	**0.80**
Supervisor	0.00	0.60	0.40	0.00	0.00	**0.60**
Peer	0.00	0.00	0.70	0.30	0.00	**0.70**
Sub-ordinate	0.00	0.00	0.40	0.60	0.00	**0.60**

A single truth-value is required for all fuzzy sets which can be arrived at by applying either an AND operator which takes the minimum value or the OR operator which takes upon the maximum value. For this appraisal system, the OR operator was selected therefore the maximum values were selected. Matlab will evaluate all rules that fire and apply the centroid method the finds the centre of gravity in the composite graph to generate the final crisp value

The defuzzification graph will be a combination of all outputs and results in an aggregate score of 50, which falls in the average category as follows:

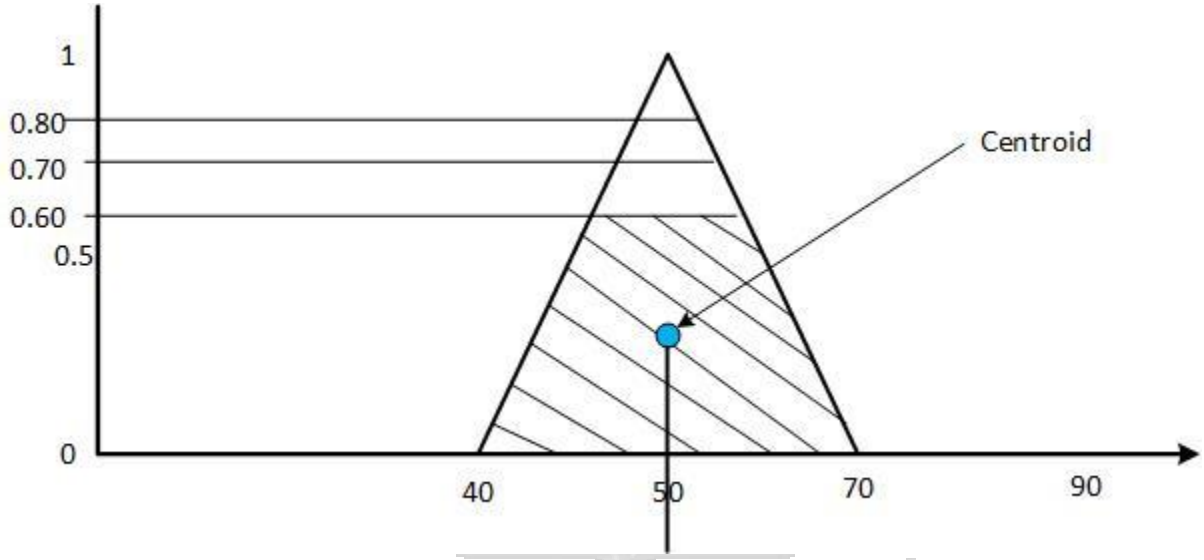

Figure 5-15Defuzzified Output

The Matlab visual of the sample-applied rules is as below:

5.5 Functional Testing

The model was tested based on the system functional and non-functional requirements to determine whether all desired functionality was achieved and to identify bugs that may result in system malfunction. The test results are summarised as follows:

Table 5.8 Results of Prototype Testing

Function	**Expected outcome**	**Results**
Login		
Valid credentials entered	Successful login	Pass
Username or Password field blank	Error reported	Pass
Wrong Username or Password entered	Error reported	Pass
User Registration		
Required Field unfilled	Error reported	Pass
Invalid data in a Field	Validate button grey	Pass

Incident Reporting		
Incomplete form	Validate button grey	Pass
View Reports		
Insufficient privileges	Error reported	Pass
Feedback		
Blank radio button	Error reported	Pass
Database Connections		
Test connection to database	Successful connection	Pass
Test update, edit, delete and read.	Operation executed	Pass

5.6 Usability Testing

Usability testing was carried out by asking the respondents involved to use the system to determine whether the system satisfies the requirements. This particular group of users are able to make reports, give feedback and view cumulative scores.

Table 5.9 Usability Testing

	Test Case	Priority	Test Results
1.	Does the system allow new user registration?	High	Successfully registered users can be viewed on the backend.
2.	Does the system successfully capture incident reports?	High	Incident reports captured and can be viewed per user.
3.	Does the system successfully capture officer field reports?	High	Field reports captured from the mobile app and stored successfully and viewed in the backend.
4.	Does the system provide real time updates to reports at the backend?	Medium	Reports from the users and officers successfully updated.
5.	Are registered users able to provide service feedback on the mobile application?	High	Users' feedback results in automatic scoring.

6.	Are supervisors able to enter annual scores for each registered officer on the backend?	High	Supervisor scores entered and stored.
7.	Does the Fuzzy Rule Base apply rules to determine Performance Score?	High	The model determined scores as per the rules.
8.	Are officers and supervisors able to view cumulative scores on the backend?	Medium	Aggregate scores are successfully displayed.

Chapter 6: Discussion

6.1 Introduction

Performance appraisal is an indispensable activity within the function of human capital management that works towards the realisation of both individual and organisational needs. This study sought to develop an appraisal model that takes into account feedback from multiple sources. Feedback from the public as a means of determining effectiveness of policing service delivery based on the level of satisfaction from members of the public who consume police services was included in the final appraisal scores. Traditional methods of quantifying an individual's performance can be subjective therefore, the resultant model is expected to enhance objectivity in the appraisal process through fuzzy logic. The study set out to review performance appraisal systems and procedures, and models used for appraisal by Police agencies. Algorithms and models used in performance appraisals were examined in order to select a suitable model that could be used to develop the appraisal model.

6.2 Existing Models Used in Appraisal in Police Agencies

The first objective was to analyse existing models used by Police Agencies in appraising the performance of their officers. The study revealed that various jurisdictions apply one or more appraisal system depending on their unique policing needs. Modern systems of appraisal such as Multi-source Feedback, Management by Objectives, Graphical Rating Scales and Critical incident reporting have been used in rating the performance of public servants. The appraisal procedures for the United Nations' Police, England and Wales, Canada, Swaziland and South Africa were studied to determine best practice and establish gaps in research. These procedures were found to be more effective with regular changes in procedure to cater for changes in policing methodologies. However, the literature revealed inadequacies in the appraisal of police officers in Kenya with recommendations to develop more objective procedures cited by agencies that include Transparency International Kenya. This justified development of the model.

6.3 Algorithms and Models used in Performance Appraisal Systems

The second objective was to analyse the algorithms and models that have been used in developing performance appraisal information systems. The literature presents the application of the Cumulated Average Weighting (CAW) algorithm, the PageRank algorithm and the Fuzzy Multi Factorial Evaluation Rule Based model in development of appraisal systems. An analysis of machine learning

algorithms along with their application in performance management was also done. Out of these algorithms, the Fuzzy Multifactorial Evaluation model was considered the most effective for the case of the model development for police appraisal. This was because of the requirement of a model void of subjectivity and bias. A model developed based on a Fuzzy Rule Base with multiple inputs as per the conceptual framework would yield consistent and accurate ratings for every officer.

6.4 Appraisal Prototype Development

The third objective was to develop an appraisal information system model. The development of the appraisal model involved use of a Fuzzy Inference System with rules that would determine a performance score output from crisp inputs. The appraisal fuzzy set contains scores from the public, supervisor, peers and subordinates as fuzzy sub-sets as per the research objectives. The fuzzy inference system with fuzzy rules was implemented at the application layer of a three-tier architecture system. The presentation layer consists of an Android based mobile front end and a web application for data input. The web-based server side performed data processing, storage and reporting. From the conceptual framework, key components of the model were outlined showing how the system components would interact to generate consistent performance scores. The public will through this model participate in determining the effectiveness of individual officers.

6.5 Appraisal Model Testing

The fourth objective was to carry out functionality and usability tests on the prototype. Validation was done by presenting users with a list of functionalities that the system must meet. The application model was found to be functional and thus addressed the research problem and passed functionality and reliability tests. The mobile and web applications required consistent input from diversified sources to effectively produce the average weighted outcome on a scale of 1 to 5. The test subjects found the system easy to use. The tools will also encourage reporting since the users will not have to visit a police station for most reporting needs as well as enabling participation in determining the effectiveness of officers.

6.6 Research Limitations

The study sought to introduce the concept of multisource feedback in the appraisal of Police officers. The application was tested and worked as per the functional requirements; however, the mobile application presents a platform constraint since it is only available for users on the Android platform.

The application relies on a stable broadband internet connection meaning users must be located where there is coverage and must have data resources.

Chapter 7: Conclusions and Recommendations

7.1 Introduction

This research sought to establish best practice in performance appraisal and in particular for police. The findings were used to develop a multisource feedback model that leverages on technology to infuse objectivity in the appraisal process of police officers. This chapter reviews the findings and achievements as well as gives recommendations for future research on the subject.

7.2 Conclusion

Performance appraisals remain a key indicator of employee competence as a means of conducting training needs analysis and other development needs. However, the process could easily turn into a formal exercise if any of the parties involved does understand its value or lacks confidence in the outcome. Rather than do away with appraisals, human capital management experts have recommended adjustments that will enrich the process including leveraging on technology to enhance objectivity in the process. Modern methods of appraisals take into account the unique operating situations of various sectors and tailor rating methods towards meeting the industry's specific needs. In the public sector, the return of investment measure in profit making organisations can be replaced with a service satisfaction index that would give taxpayers an indication of quality in service delivery.

The study sought to understand existing appraisal systems used by law enforcement agencies and to identify gaps that would inform development of an objective approach for appraising police officers in Kenya. The research revealed inadequacy in present systems in Kenya that essentially place the appraisal process at the discretion of supervising officers who leave room for subjectivity in scoring due to personal bias or preferences. This has been shown as a major cause of demoralisation since appraisals are relied upon in administrative procedures such as determining promotions and other rewards. A critical review of performance appraisal systems revealed that this anomaly could be resolved by sourcing appraisal data from multiple sources. An important source of this data in jurisdictions where policing is a service rather than a force is feedback from the consumers of policing services - members of the public. Police reforms in Kenya are geared towards establishing policing as a service that can be trusted by members of the public and this would be achieved by enabling the consumers of policing services to have input in determining the effectiveness and quality of service delivered. An objective process will restore relevance to the appraisal process, which has largely been

a formality to comply with regulation and compel officers to improve service delivery to their customers.

This study presented a model that satisfied the identified requirements key to an objective appraisal system and its implementation would generally improve Police appraisal systems and motivate officers to work harder to earn good scores and public trust. The system adopted a three-tier architecture and uses a fuzzy inference system, which computes and gives uniform scores to all officers based on standard previously established evaluation criteria (fuzzy rule base). The mobile interface enables feedback from the public following reporting of an incidence and resolution of the same by an assigned officer. This feedback along with ratings from an officer's supervisor and partners complete a 360-degree appraisal, which gives an indication of performance from all parties that interact with the officer while performing their duties. The use of a rule base eliminates errors in computation and establishes uniformity in scoring which gives objective feedback for the officer on their performance and informs their superiors on their developmental needs. This study has demonstrated that the appraisal process of police officers can gain greater acceptance by all parties involved by leveraging on technology.

7.3 Recommendations

The study gives recommendations as follows:

i. The study recommends adoption of information technology in Multisource feedback for performance appraisal of all public servants where a return on investment can be measured by quality in service delivery to taxpayers. An understanding that performance feedback will be sourced from those who interact with an employee would not only lead to improved service delivery but also promote healthy workplace practices in the National Government, County Governments, State Corporations and Agencies.

ii. The study recommends continuous sensitisation to all employees in the public sector on the importance of performance appraisals as both a measure of competence and a means to determine employee development needs. The sensitisation along with training on conducting the appraisals will ensure that the importance of the process will be appreciated and adoption of technology accepted therefore validating appraisals not just another

compliance exercise but also one that would yield quality results that can be relied upon when executing human capital management decisions.

7.4 Future work

i. The Fuzzy Rule Base can be created using the self-learning algorithm (Neuro-Fuzzy System) where part of the input data will be used to train the system and construct the rule base.

ii. The mobile-based client front end can be improved by developing a cross-platform mobile application to cater for users who prefer other operating systems other than Android. The system could also be further developed to have USSD capability to cater for users who do not have smartphones.

iii. To maintain confidentiality of individual appraisal scores, backend access could be restricted to a particular range of IP addresses to restrict access that will ensure improved security levels.

References

Aggarwal, A., & Thakur, G. S. M. (2013). Techniques of performance appraisal-a review. International Journal of Engineering and Advanced Technology (IJEAT), 2(3), 617-621.

Agarwal, P., Hunt, K., Zhuang, J., Sarkar, B., Sarkar, A., & Sharma, R. (2019). An exploratory analysis for performance assessment of state police forces in India: an eclectic approach. *Operational Research*, 1-27.

Ahsan, M. Z., Khandkar, M. R., & Islam, M. S. (2016). Algorithm for performance appraisal using Cumulative Average Weighing Method.

Alzubi, J., Nayyar, A., & Kumar, A. (2018, November). Machine learning from theory to algorithms: an overview. In Journal of physics: conference series (Vol. 1142, No. 1, p. 012012). IOP Publishing.

Amra, I. A. A., & Maghari, A. Y. (2017). Students performance prediction using KNN and Naïve Bayesian. In 2017 8th International Conference on Information Technology (ICIT) (pp. 909-913). IEEE.

Anstey, E., Riggar, T. F., & Walker, J. (2017). *Staff appraisal and development*. Routledge.

Ayers, R. S. (2015). Aligning individual and organizational performance: Goal alignment in federal government agency performance appraisal programs. *Public Personnel Management*, *44*(2), 169-191.

Bagayoko, N., Hutchful, E., & Luckham, R. (2016). Hybrid security governance in Africa: rethinking the foundations of security, justice and legitimate public authority. *Conflict, Security & Development*, *16*(1), 1-32.

Barasa, S. W. (2017). Influence of Work Environment on Performance in the Public Security Sector with a Focus on the Police in Nairobi Kenya. OIDA International Journal of Sustainable Development, 10(11), 49-64.

Barrett, B. J., Peirone, A., & Cheung, C. H. (2016). Spousal Violence and Evaluations of Police Performance in Canada: Does Police Contact Matter?. *Journal of Family Violence*, 1-13.

Batts, A. W., Smoot, S. M., & Scrivner, E. M. (2012). Police leadership challenges in a changing world. US Department of Justice, Office of Justice Programs, National Institute of Justice.

Belle, N., Cantarelli, P., & Belardinelli, P. (2017). Cognitive biases in performance appraisal: Experimental evidence on anchoring and halo effects with public sector managers and employees. *Review of Public Personnel Administration*, *37*(3), 275-294.

Bennett, T. (2019). *Improving public sector performance through standardized human capital performance measurement & reporting* (Doctoral dissertation).

Berger, P. L. (Ed.). (2019). *A future South Africa: visions, strategies, and realities*. Routledge.

Bianchi, C. (2010). Improving performance and fostering accountability in the public sector through system dynamics modelling: from an 'external' to an 'internal' perspective. *System Research & Behavioural Science,* 27(4), p. 361–384.

Black, S., Gardner, D. G., Pierce, J. L., & Steers, R. (2019). Techniques of Performance Appraisal. Organizational Behavior.

Bridgelall, R. (2017). Introduction to Support Vector Machines. *Lecture Notes*, 1-18.

Chan, D. C. (2006). Core competencies and performance management in Canadian public libraries. Library management.

Chitroda, N. K., & Katurbadham, T. (2014). Improve Page Rank algorithm using normalized technique. International journal of data warehousing & mining, 4(1).

Chung, E. K., Jung, Y., & Sohn, Y. W. (2017). A moderated mediation model of job stress, job satisfaction, and turnover intention for airport security screeners. Safety science, 98, 89-97.

Cintrón, R., & Flaniken, F. (2011). Performance appraisal: A supervision or leadership tool. International Journal of Business and Social Science, 2(17), 29-37.

Chappell, D., & Evans, J. (2008). The role, preparation and performance of civilian police in United Nations peacekeeping operations. *Commonwealth Law Bulletin*, *24*(3-4), 1248-1316.

Ding, J., Zheng, Q., Wang, X., Zhu, H., & Zhang, J. (2016). Assessment of Innovative Performance Management in Chinese Police System: A Case Study of Zhejiang Province. *Public Personnel Management*, *45*(1), 6-25.

Elliott, J., & Lal, S. (2016). Blood pressure, sleep quality and fatigue in shift working police officers: effects of a twelve hour roster system on cardiovascular and sleep health. *International journal of environmental research and public health*, *13*(2), 172.

Ermongkonchai, P. (2008). An evaluation of multisource feedback (MSF) for managerial development in large-size manufacturing companies in Thailand. Contemporary Management Research, 4(3).

Faull, A. (2016). Measured governance? Policing and performance management in South Africa. *Public Administration and Development*, *36*(2), 157-168.

Gallab, M., Bouloiz, H., Alaoui, Y. L., & Tkiouat, M. (2019). Risk assessment of maintenance activities using fuzzy logic. *Procedia computer science*, *148*, 226-235.

Gao, X., Wen, J., & Zhang, C. (2019). An improved random forest algorithm for predicting employee turnover. Mathematical Problems in Engineering, 2019.

Good, D., & Coombe, D. (2009). Giving multisource feedback a facelift. Journal of Change Management, 9(1), 109-126.

Gibbons, F. X., & Kleiner, B. H. (2013). Factors that bias employee performance appraisals. *Management Research News*.

Graham, R., & Beuthin, R. (2018). Exploring the effectiveness of multisource feedback and coaching with nurse practitioners. Nurs Leadersh (Tor Ont), 31(1), 50-59.

Hassan, M. U., Shaukat, S., Nawaz, M. S., & Naz, S. (2013). Effects of innovation types on firm performance: an empirical study on Pakistan's manufacturing sector. *Pakistan Journal of Commerce and Social Sciences*, *7*(2), 243-262.

Himanshu, K. S., Kshitij, V., & Srinivasan, R. (2020). Employee performance and leave management using data mining technique. *Journal of Pure and Applied Mathematics.*

Hoffer, J. A. (2012). *Modern Systems Analysis and Design, 6/e*. Pearson Education India.

Hope, K. R. (2016). An analytical perspective on police corruption and police reforms in developing societies. *Police corruption and police reforms in developing societies*, 3-31.

Ismail, H. N., & Gali, N. (2017). Relationships among performance appraisal satisfaction, work–family conflict and job stress. *Journal of Management & Organization*, *23*(3), 356-372.

Islami, X., Mulolli, E., & Mustafa, N. (2018). Using Management by Objectives as a performance appraisal tool for employee satisfaction. *Future Business Journal*, *4*(1), 94-108.

Jacobs, G., Belschak, F. D., & Den Hartog, D. N. (2014). (Un) ethical behavior and performance appraisal: the role of affect, support, and organizational justice. *Journal of business ethics*, *121*(1), 63-76.

Jahan, S. (2014). Human resources information system (HRIS): A theoretical perspective. *Journal of Human Resource and Sustainability Studies*, 2, 33-39.

Janssens, L., Smet, K., Onghena, P., & Kyndt, E. (2017). The relationship between learning conditions in the workplace and informal learning outcomes: a study among police inspectors. *International Journal of Training and Development*, *21*(2), 92-112.

Jantan, H., Yusoff, N. M., & Noh, M. R. (2014). Towards Applying Support Vector Machine Algorithm in Employee Achievement Classification. In International Conference on Data Mining, Internet Computing, and Big Data.

Jayadi, R., Jayadi, R., & Firmantyo, H. M. (2019). Employee Performance Prediction using Naïve Bayes. In *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (Vol. 8, No. 6, pp. 8-12).

Jin, C., De-Lin, L., & Fen-Xiang, M. (2009). An improved ID3 decision tree algorithm. In *2009 4th International Conference on Computer Science & Education* (pp. 127-130). IEEE.

Jones, C. N. (2018). Human Resources Professionals Perceptions of Gamifying Performance Appraisals (Doctoral dissertation, Emporia State University).

Kamińska, B., & Szymańczyk, K. (2014). Professional Evaluation. Application of the PageRank Algorithm in Employee Rating. Przedsiebiorczosc i Zarzadzanie, 15(2), 37-50.

Kakulapati, V., Chaitanya, K. K., Chaitanya, K. V. G., & Akshay, P. (2020). Predictive analytics of HR-A machine learning approach. Journal of Statistics and Management Systems, 23(6), 959-969.

Kwon, Y., Balazinska, M., Howe, B., & Rolia, J. (2012). Skewtune: mitigating skew in mapreduce applications. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data* (pp. 25-36). ACM.

Landy, F., Zedeck, S., & Cleveland, J. (2017). Performance measurement and theory. Routledge.

Mabila, N. O. (2014). Job satisfaction in the Royal Swaziland Police Service*: a case study of Manzini and Hhohho regions* (Doctoral dissertation).

Magesh, N., Thangaraj, P., Sivagobika, S., Praba, S., & Priya, R. M. (2013). Evaluating The Performance Of An Employee Using Decision Tree Algorithm. Int. J. Eng. Res. Technol, 2(4), 2814-2830.

Mahesh, B. (2020). Machine Learning Algorithms-A Review. International Journal of Science and Research (IJSR).[Internet], 9, 381-386.

Majid, J. (2016). Effectiveness of Performance Appraisal Methods–An Empirical Study of the Telecommunication Sector. International Journal of Trend in Research and Development, 3(3), 10-17.

Makabira, D. K., & Waiganjo, E. (2014). Role of Procurement Practices on the Performance of Corporate Organizations in Kenya: A case study of Kenya National Police Service. *International Journal of Academic Research in Business and Social Sciences*, *4*(10), 369.

Meenakshi, G. (2012). Multi source feedback based performance appraisal system using Fuzzy logic decision support system. *arXiv preprint arXiv:1203.1882.*

Metcalf, B. (2017). New police management, performance and accountability. In *Questioning the New Public Management* (pp. 71-88). Routledge.

Mu, L. (2016). Research on Police Professional Quality Training Methods of Police Academy Students. In *2016 2nd International Conference on Education Technology, Management and Humanities Science*. Atlantis Press.

Murphy, K., Sargeant, E., & Cherney, A. (2015). The importance of procedural justice and police performance in shaping intentions to cooperate with the police: Does social identity matter? *European journal of criminology*, *12*(6), 719-738.

Nair, M. S., & Salleh, R. (2015). Linking performance appraisal justice, trust, and employee engagement: A conceptual framework. *Procedia-Social and Behavioral Sciences*, *211*, 1155-1162.

Nguyen, L. (2017). Tutorial on support vector machine. Applied and Computational Mathematics, 6(4-1), 1-15.

O'Toole Jr, L. J., & Meier, K. J. (2014). Public management, context, and performance: In quest of a more general theory. *Journal of public administration research and theory*, *25*(1), 237-256.

Ozkan, C., Keskin, G. A., & Omurca, S. I. (2014). A variant perspective to performance appraisal system: Fuzzy C–Means algorithm. *International Journal of Industrial Engineering, 21(3).*

Pasha, O. Q. (2014). Performance management system design and implementation in police agencies: Is following recommended practices worth it? (Doctoral dissertation, Georgia Institute of Technology).

Paul, D., Agarwal, P., & Chakraborty, S. (2016). Performance appraisal of Indian state police forces using ARAS method. *Management Science Letters*, *6*(5), 361-372.

Pälli, P., & Lehtinen, E. (2014). Making objectives common in performance appraisal interviews. *Language & Communication*, *39*, 92-108.

Perez-Floriano, L. R., & Gonzalez, J. A. (2019). When the going gets tough: A moderated mediated model of injury, job-related risks, stress, and police performance. *International Journal of Organizational Analysis, 9(*4), 242-259.

Phin, L. W. (2015). The Effectiveness of Performance Appraisal in the Private Education Industry in Malaysia. *International Journal of Business & Information*, *10*(1).

Pollitt, C. (2013). The logics of performance management. *Evaluation*, *19*(4), 346-363.

PSC 37A (2016). Public service commission. Staff performance appraisal report.

Ransley, P. (2009). Republic of Kenya; *Report of the national task force on police reforms.*

Rand centre on quality policing. Technical report (2012).

Rao, T. V. (2004). Performance Management and Appraisal Systems: HR tools for global competitiveness. SAGE Publications India.

Saunders, M., Lewis, P. & Thornhill, A., (2013). *Research methods for business students.* 6th Edition. Cape Town: Pearson Education.

Saxena, R. (2010). Role of Communication in Employee Development. International Transactions in Humanities & Social Sciences, 2(2).

Schultz, J. R. (2015). To improve performance, replace annual assessment with ongoing feedback. *Global Business and Organizational Excellence*, *34*(5), 13-20.

Shah, A. D., & Ladhake, S. A. (2014). Multi user feedback system based on performance and appraisal using Fuzzy Logic Base system-design and implementation. *International Journal of Engineering Research and Application, 674-680.*

Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. IEEE Access, 8, 80716-80727.

Siraj, N. B., & Fayek, A. R. (2019). Risk identification and common risks in construction: Literature review and content analysis. Journal of Construction Engineering and Management, 145(9), 03119004.

Suh, Y. A., & Kim, J. (2020). Estimation of the likelihood of severe accident management decision-making using a fuzzy logic model. Annals of Nuclear Energy, 144, 107581.

Transparency International Kenya (2016). Police service satisfaction survey and needs analysis report. A focus on Kisumu and Nairobi Counties.

Tziner, A., & Rabenu, E. (2018). Ways to improve the performance appraisal system 2: Alternative strategies for assessing and evaluating performance. In *Improving Performance Appraisal at Work*. Edward Elgar Publishing.

Van Thielen, T., Bauwens, R., Audenaert, M., Van Waeyenberg, T., & Decramer, A. (2018). How to foster the well-being of police officers: The role of the employee performance management system. *Evaluation and program planning*, *70*, 90-98.

Walker, S. (2015). Introduction to early intervention systems. *Critical Issues in Policing: Contemporary Readings*, 244.

Werner, J. M., & DeSimone, R. L. (2011). Human resource development. Cengage Learning.

Williams, B. N., Christensen, R. K., LePere-Schloop, M., & Silk, P. D. (2015). Appraising the appraisal process: Manager and patrol officer perspectives. *The Police Journal*, *88*(3), 231-250.

Yang, K., & Holzer, M. (2006). The performance–trust link: Implications for performance measurement. *Public Administration Review*, *66*(1), 114-126.

Yee, C. C., & Chen, Y. Y. (2009). Performance appraisal system using multifactorial evaluation model. *World Academy of Science, Engineering and Technology, 53(2009), 231-235.*

Appendix A: Sample Code (Algorithm Score Computation)

The Source code snippet for aggregation of scores through the CAW Algorithm is as below:

```
public function variance()
{
$police = $this->Users->get_policemen();
foreach($police as $p)
{

$re = $this->Users->get_feedback_scores($p['idno']);
$array_sum = 0;
$the_variance = 0.0;
$the_mean = 0.0;

$no_elements= count($re);
foreach($re as $row)
{
$array_sum = $array_sum + $row['feedback_score'];
}

$the_mean = $array_sum / $no_elements;

foreach($re as $row)
{
$the_variance = $the_variance + ($row['feedback_score']- $the_mean) * ($row['feedback_score'] -
$the_mean);
}
$the_variance = $the_variance / $no_elements;
$this->Users->record_scores($p['idno'],$the_variance);
}

$this->session->set_flashdata('successg', '<p  class="alert alert-success">Police Scores
Successfully Generated, check on the reports menu</p>');

redirect('welcome/generate');
}
```

Appendix B: Mobile Application Screenshots

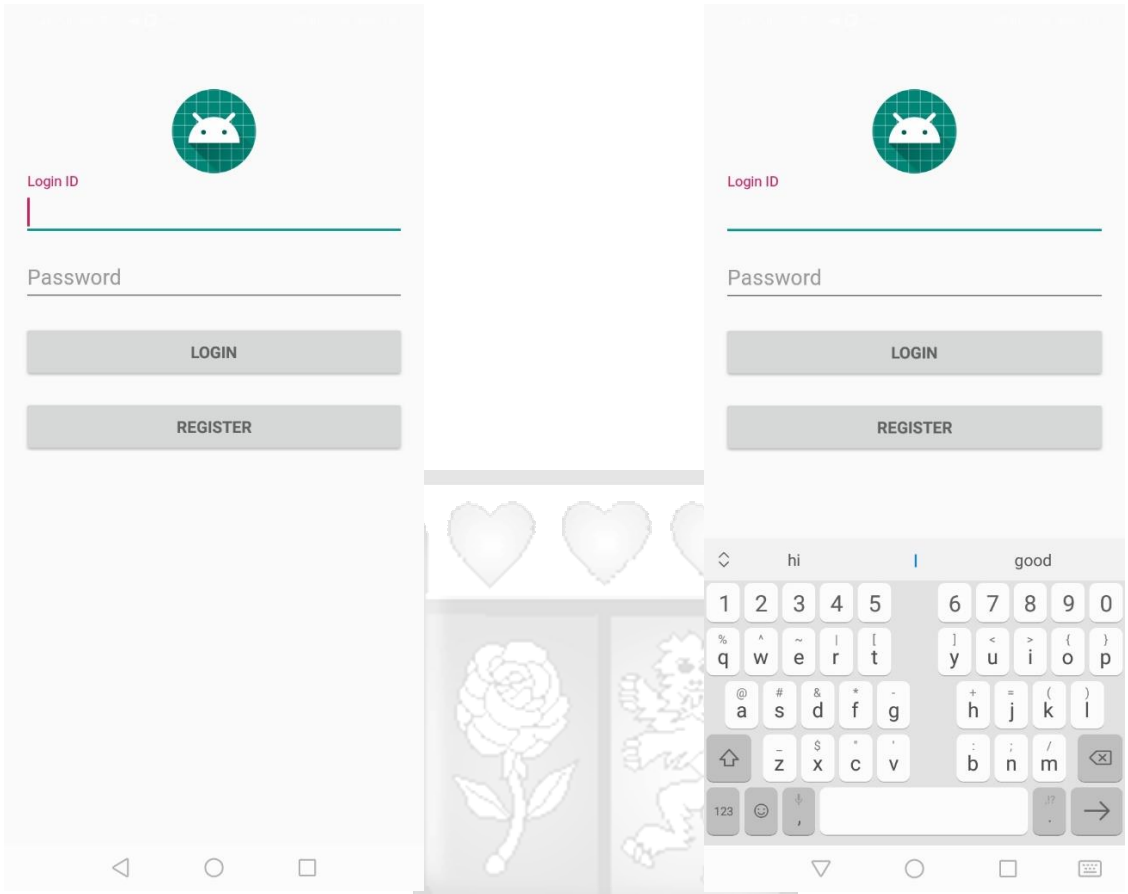

Figure C1 Login Screen

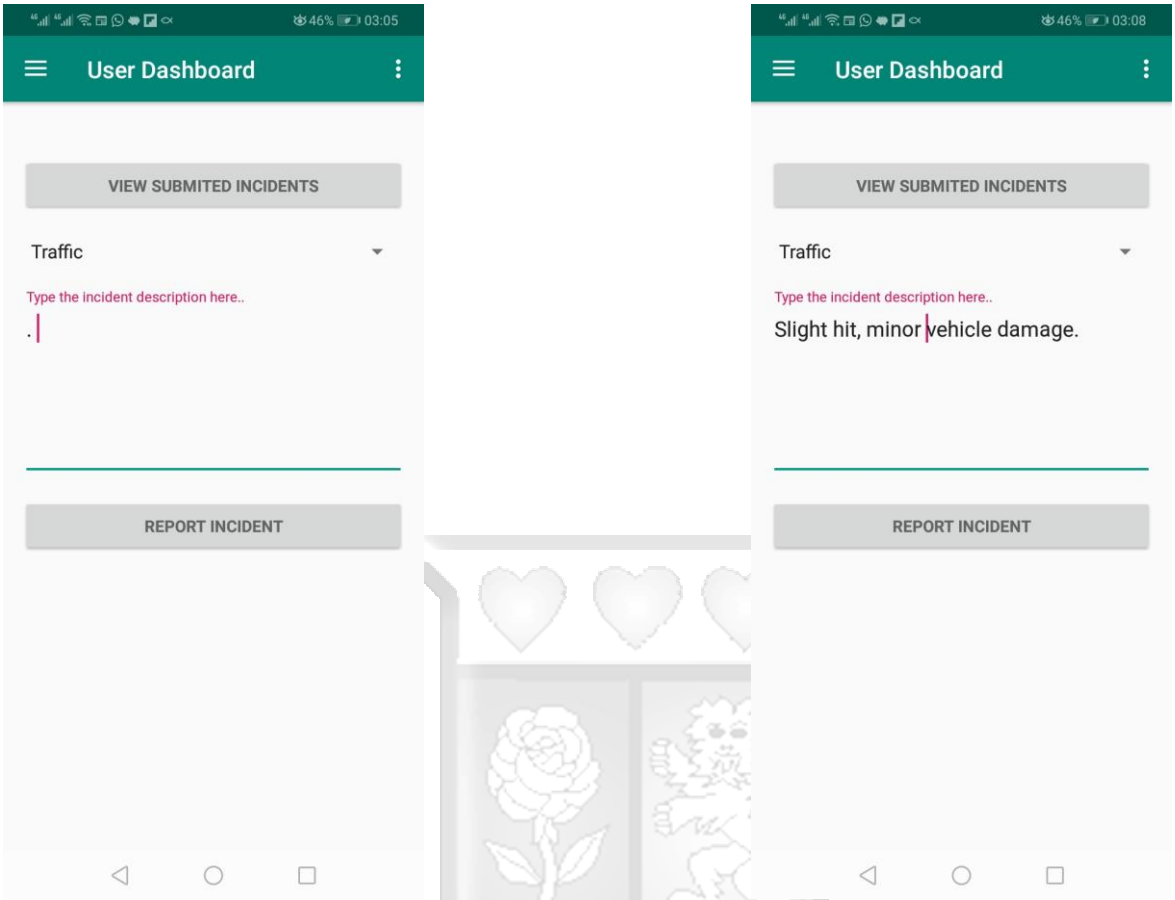

Figure C2 Public User Report Screen

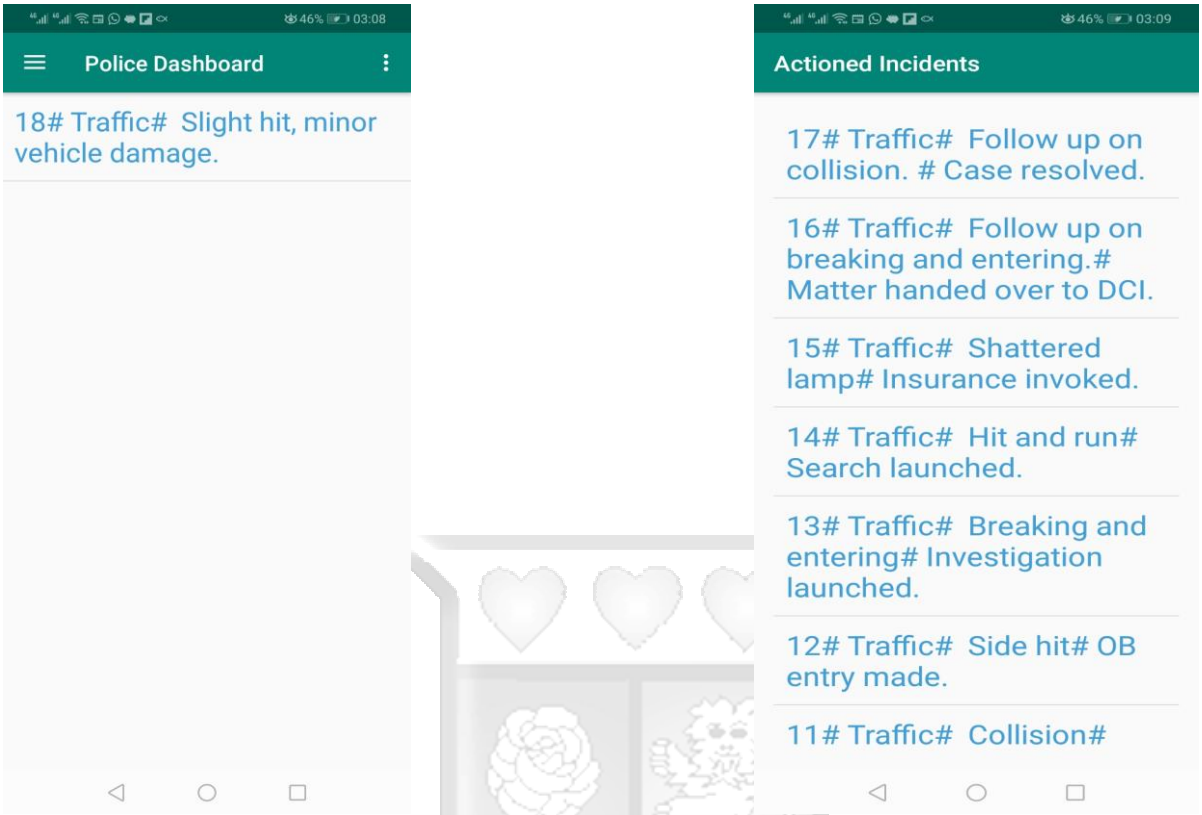

Figure C3 Field Officer Action Screens

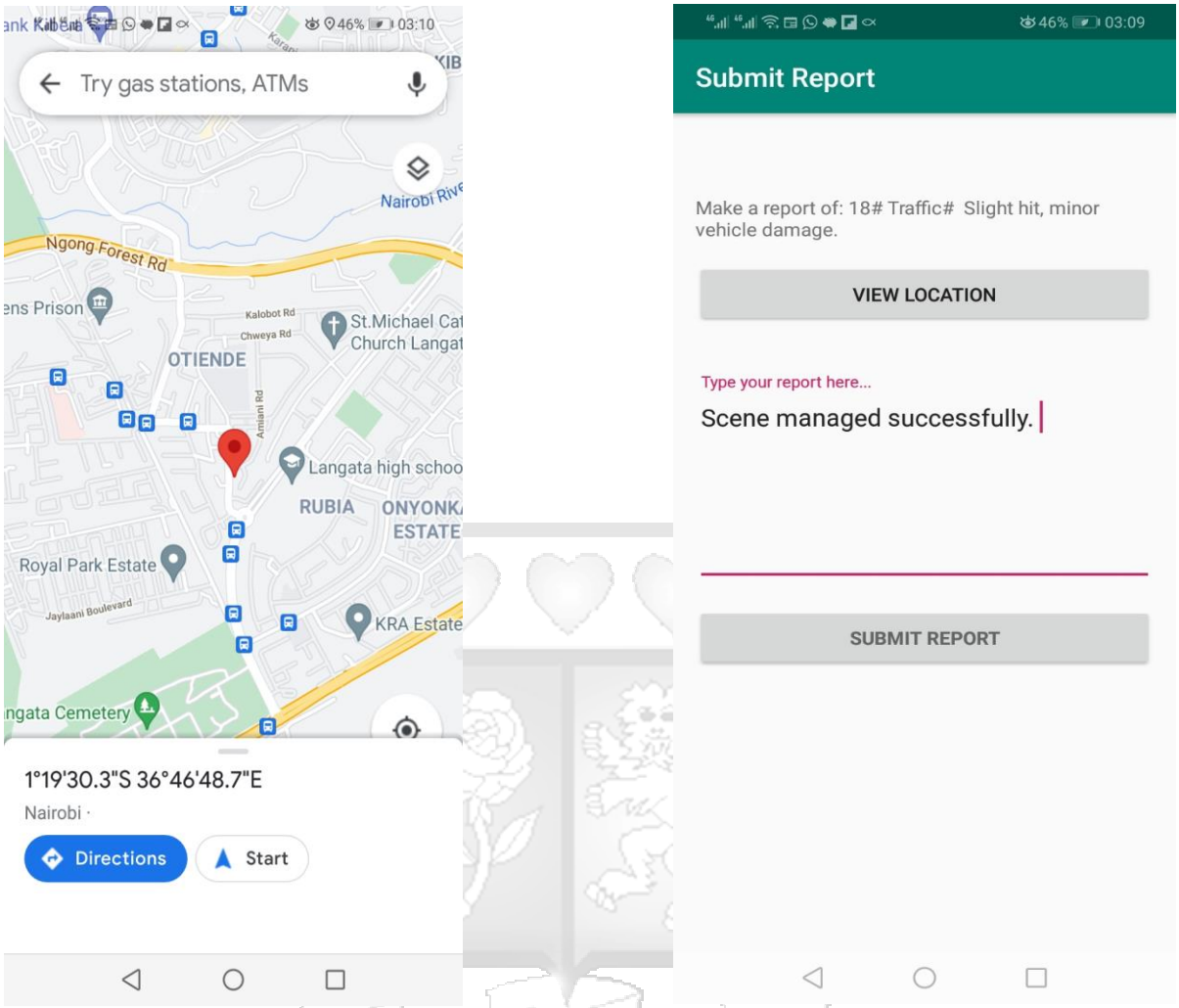

Figure C4 Field Officer Incident Locator Map and Scene Report Upload Screen

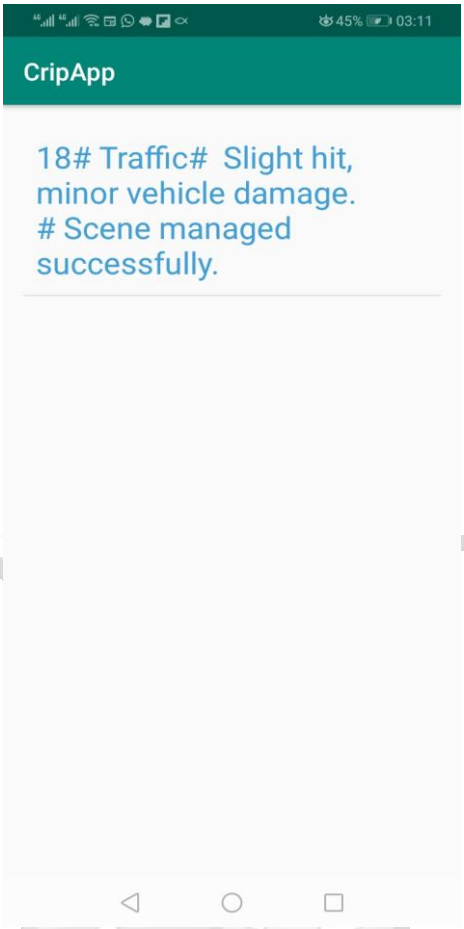

Figure C5 Public User Actioned Incident Page

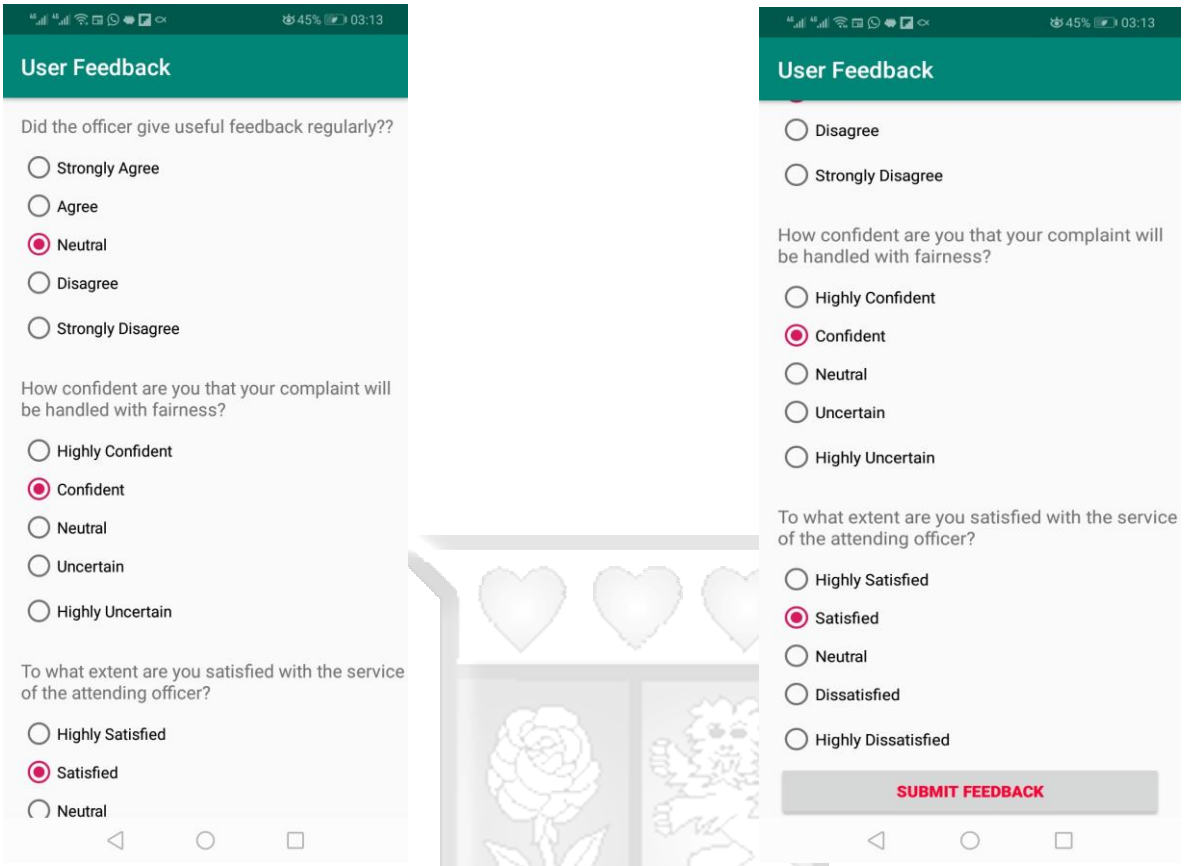

Figure C6 Public User Feedback Form