Applying data analytics for enhanced construction project performance through structural concrete rework predictive models.

by

Fionah Mazvita Mukondwa

Thesis presented in fulfilment of the requirements for the degree of Master of Engineering in Civil Engineering in the Faculty of Engineering at Stellenbosch University

Supervisor: Professor Jan Andries Wium

December 2021

DECLARATION

I Fionah Mazvita Mukondwa declare that the contents of this thesis are a presentation of my original research. Wherever contributions of other researchers are involved, efforts were made to clearly indicate these instances, with due reference and acknowledgement to the relevant literature.

The research was carried out under the guidance and supervision of Professor Jan Andries Wium, at the University of Stellenbosch, South Africa.

…………………………………….. …………………………………………

Fionah Mazvita Mukondwa Date

In my capacity as academic supervisor of the above-mentioned candidate's thesis, I acknowledge and certify to the best of my knowledge, that the aforementioned declaration is indeed, true.

…………………………………….. …………………………………………

Professor J.A Wium Date

ABSTRACT

Today's world is driven by data-based decision-making that needs to be accurate to effectively solve engineering problems involving the prediction of failure, defects, and errors. Motivated by the fourth industrial revolution (Industry 4.0) that has enhanced the performance of construction industries on a global scale, this study discusses the development of a predictive machine learning model that can be used during the construction phase to manage structural concrete rework during site inspections. This model seeks to reduce uncertainties and minimise structural concrete rework during construction to enhance project performance. To develop the model, the research approach included an exploratory case study together with interviews with experienced professionals in structural concrete construction at the Hwange Expansion Project, a mega thermal power plant construction project in Hwange, Zimbabwe. The exploratory case study and expert interviews were conducted to establish a better understanding of the risk triggers that influence structural concrete rework in a typical construction project.

A fictitious modelling dataset was then generated based on the results of a questionnaire survey conducted on structural concrete experts due to the lack of sufficient project data. Various data mining techniques were also employed to develop the prediction model following some steps of the Cross Industry Standard Practise for Data Mining (CRISP-DM) framework. This fictitious dataset was modelled on five classification algorithms whose performance was evaluated using the 20-fold cross-validation test. The Neural Network classifier recorded the highest performance with accuracy and precision of over 95%. To validate the performance of the Neural Network prediction model, the confusion matrix validation test was carried out on six datasets of varying size ranging from 500 to 10 000 data points. The results from the confusion matrix validation test indicated, as expected, that the larger the dataset, the more accurate and robust the model becomes in predicting new data outcomes.

Based on these findings, it was established that data analytics in the form of predictive modelling can be used by organisations to reduce uncertainties and promote data-driven decision-making during structural concrete quality checks on site. It is recommended that construction industries employ data analytics as a project management tool not only to enhance the performance of construction projects but to build reference databases for further development of big data in the industry.

ACKNOWLEDGEMENTS

Firstly, I acknowledge and appreciate the invaluable guidance my academic supervisor, Professor J.A. Wium offered me through the course of my research. Regardless of the negative impacts the Covid-19 pandemic had on the progress of my studies, my supervisor's tireless efforts ensured the successful completion of the research.

Secondly, I extend my profound gratitude to my good friends Bernard Pungu, Nyasha Zamimmba, and Silas Kashiri who remained supportive and offered steadfast moral support to ensure my research efforts become a success. Special mention goes to my colleagues and superiors at the Hwange Expansion Project, whose unparalleled professional input and guidance immensely contributed to the development of the study.

Deep appreciation also goes to my parents, Fidelia and Joseph Mukondwa whos' unconditional love and support gave me strength whenever the going got tough, and in their own unique way, made my research journey all the more worthwhile.

Lastly, to the Lord, my God Almighty, I extend deep thanksgiving for the protection, wisdom, and guidance He bestowed upon me during the undertaking of my research.

ABBREVIATIONS

CRISP-DM	Cross Industry Standard Process for Data Mining
EPC	Engineering Procurement and Construction
FIDIC	French acronym interpreted in English as International Federation of Consulting Engineers
GDP	Gross Domestic Product
HEP	Hwange Expansion Project
HIRA	Hazard Identification and Risk Assessment
HTPS	Hwange Thermal Power Station
ISO	International Organisation of Standardisation
ML	Machine Learning
MW	Megawatt
NCR	Non-Conformance Report
PMBOK	Project Management Body of Knowledge
PRINCE 2	Projects in Controlled Environment
QA	Quality Assurance
QC	Quality Control
QMP	Quality Management Plan
SDG	Sustainable Development Goal
SU	Stellenbosch University
USD	United States Dollar
UN	United Nations
ZIE	Zimbabwe Institution of Engineers
ZPC	Zimbabwe Power Company

DEDICATION

To my grandparents Helen and Wenceslaus Mkumbuzi

You give me purpose

Table of Contents

List of Figures

List of Tables

CHAPTER ONE

RESEARCH INTRODUCTION

1.1 Introduction

This chapter introduces the reader to the primary focus of the research which is investigating the feasibility of implementing data analytics for enhanced construction project performance through structural concrete rework predictive models. Presented in various sections, the chapter outlines the problem statement enabling the reader to appreciate the major issues surrounding the motivation of the study. The research is motivated by the persistence of construction field rework occurrence that continues unabated throughout construction industries although several researchers have developed rework management and containment strategies.

Additionally, Industry 4.0 has revolutionised the construction industries of developed countries and this research seeks to promote digitalization and data-driven decision-making in the Zimbabwean construction industry. To achieve this, the study investigates the various structural concrete rework risk triggers that occur in a typical construction project during the construction phase. Descriptive statistical methods are used to analyse the collected data to compile a dataset that can be used in the development of a predictive model that can manage construction field rework during the project construction phase.

1.2 Research background

The economic development and growth of any country are influenced by the performance of the various economic sectors that contribute to the national Gross Domestic Product (GDP). One of the sectors that significantly stimulates economic development is the construction sector. Activities involved in construction especially in developing countries influence the accomplishment of various socio-economic developmental goals such as infrastructure development and employment creation (Oladinrin et al., 2012). In addition to that, construction activities empower communities through the accomplishment of some of the 17 Sustainable Development Goals (SDGs) formulated by the United Nations in 2017 which support *"Promoting sustainable economic growth and full employment"* and "*Building resilient infrastructure, promoting inclusive and sustainable industrialization and fostering innovation*." to mention but a few.

In Zimbabwe, the construction industry contributed almost 3% to the national GDP for the year 2016 according to ZimStat (2016) with growth projections aimed at 10% (of GDP) by 2020. The 3% contribution recorded in 2016 is relatively low compared to Ghana, Nigeria, Malaysia, and China whose construction sectors contributed an average of 8.95%, 5%, 5%, and 5.7% respectively (World Economic Outlook, 2016). The reason behind this low input may be attributed to several factors that influence the overall success of construction projects in Zimbabwe.

Successful construction project completion within the specified budget and timeframe has become a common challenge throughout the international community. It is uncommon that a project is executed and completed well within the stipulated schedule, budget, and desired product quality based on the schedule and cost overruns that are experienced in construction projects on a global scale (Love et al., 2016). Project managers have the responsibility of ensuring a healthy balance between competing demands of quality, scope, schedule and cost (PMI, 2010). Good construction project management is therefore critically important in the successful completion of projects.

An example of good management practices adopted over the years is the introduction of construction technology promulgated by the 4th industrial revolution. Industry 4.0 characterized by the assimilation of the digital, biological, and physical domains and increase in the use of new technologies such as artificial intelligence and cloud computing among others, has instigated enhanced project management in the construction sector. This has necessitated the use of Unmanned Aerial Vehicles (drones) for site surveillance, artificial intelligence techniques in the prediction of cost overruns for budgeting and planning purposes as well automated and robotic equipment.

The construction industry is slowly leaning towards automation and digitalization relying less on human judgment for decision-making. Although technological advancements have slowly infiltrated the industry, it has not completely rid the industry of human interaction. Therefore, errors or mistakes are bound to occur. These human “errors” usually lead to rework which can occur any time throughout the construction project cycle, significantly affecting the performance of the project in terms of schedule, cost, and product quality.

Rework is a by-product of human-error and inefficient/inappropriate quality management strategies that can occur during the construction project cycle. Approximately 50% of errors in building construction originate from the design stage whilst 40% emanate during project execution/construction (Love et al., 2016). Similarly, Love (2019) reports that although design-related errors significantly contribute to rework-related costs, the majority of rework incidents usually occur during the construction phase increasing indirect project costs due to substantial construction delays and compromised product/service quality.

Construction errors emanate from active and latent errors that usually remain undetected during construction. Anyone can make these errors regardless of the level of education, competency, and experience according to Love et al. (2011). Construction errors as defined by Rosowsky and Stewart (1996) are deviances in the original design implementation because of human related errors, assuming the design is in full compliance with the stipulated design standards and codes of practice. Busby (2001) is also of the opinion that errors are bound to arise during construction when site operatives fail to effectively carry out organizational mechanisms. Construction errors are therefore still prevalent because of errors in human judgement according to Yates and Lockley (2002).

1.3 Problem statement

The performance of construction projects is significantly affected by construction rework, which according to research, results in schedule and budget overruns as well as compromised quality and structural integrity (Hwang et al., 2009). Construction-related rework can account for up to 20% of construction costs during the construction phase (Manuel et al., 2018). Additionally, there are no proactive measures put in place to enable construction professionals and individual organisations to forecast the probability of rework in structural concrete elements (Safapour and Kermanshachi, 2019) in order to effectively minimise and manage it.

1.4 Motivation of research

Firstly, this research was motivated by the persistence of construction field rework occurrence that continues unabated although several researchers have developed innovative rework management and containment strategies. Numerous studies have contextually explored the various risk triggers that lead to rework occurrence, but limited studies have suggested strategies to proactively eliminate the occurrence of structural concrete rework in particular (Love et al., 2016).

Furthermore, there are no standard measures to identify, record, track and monitor rework occurrence in the Zimbabwean construction industry, making it difficult to implement effective and suitable prevention and management of rework strategies according to personal communications with a Zimbabwean based project manager with twelve years of construction project management experience.

For these reasons, this research aims at enhancing the existing knowledge and developing a model that can be used as a project management tool to predict rework occurrence during construction. Also, since Industry 4.0 has necessitated data-based decision making, this research aims to use this as a proactive measure to minimize and eliminate structural concrete rework in the Zimbabwean construction industry.

1.5 Justification of research

As highlighted by Love (2001) the majority of rework incidents usually occur during construction hence the need for robust rework minimization and management strategies during project implementation. Similarly, Love et al. (1999) and Love and Li (2000) conducted a series of investigations and came to the conclusion that for industrial construction projects, rework emanates from construction change errors. Unlike design-related rework, construction rework involves physical activities which require project resources such as adding more reinforcement to a concrete beam which is costly and time-consuming. It has been argued that design rework constitutes a higher fraction of overall project rework hence various proactive design rework reduction and management strategies have been formulated over the years (Love et al., 2004) and (Arundachawat et al., 2009).

Love and Sing (2011) suggest that for the effective management of construction rework, it is important to determine the probability of its occurrence since the good foresight of rework risk triggers can help minimize them. As such, several tools have been developed in an attempt to reduce rework occurrences (Zhang 2009; Mastenbroek, 2010; Basbeth, 2014), however, there is a need for further analysis and improvement of these techniques since the problem has continued unabated. The predictive model in this research is based on supervised machine learning algorithms and seeks to bridge the gap in knowledge between rework elimination/reduction and predictive data analytics in the construction industry as far as structural concrete elements are concerned. Machine learning is an emerging field in the predictive analysis of data as it integrates data-based analytics, data mining, statistics, and pattern recognition to realise interrelationships and trends from various datasets.

1.6 Aim and objectives of research

This study aims to investigate the feasibility of developing a predictive model that can be used during the project implementation phase to minimize and manage structural concrete rework on site. To achieve this, the study consisted of three parts, the identification of structural concrete rework risk triggers as well as the capturing and analysis of rework data on a specific construction project. To achieve the aim of the research, the following objectives were developed:

1. To identify construction risk triggers that lead to structural concrete rework during the project construction phase.
2. To categorize the identified risk triggers into phases during reinforced concrete structure construction.
3. To develop a suitable dataset that can be used to train and test the performance of classification algorithms.
4. To apply predictive data analytics in developing a model that can be used to manage the occurrence of construction field rework.

1.7 Research design

The research used a hybrid design approach to develop the machine learning model. Qualitative research methods such as physical site inspections, semi-structured interviews with various site personnel, and analysis of site documentary sources were considered in obtaining data on current construction errors that lead to rework. Hwange Expansion Project a mega Engineering, Procurement, and Construction (EPC) project which encompasses the construction of two 335MW coal-fired generating units was investigated. The data from the exploratory study was captured in a Microsoft Excel spreadsheet which was processed to populate the model.

1.8 Research scope

The research scope mainly focused on structural concrete rework where elements such as beams, columns, and slabs were cast. Additionally, the study focused on the rework resulting from poor quality management and construction-related errors during the project implementation phase rather than design-related errors that arise during the design period. The project was carried out by contractors and subcontractors certified by the International

Organisation of Standardization (ISO) 9001;2015 who were mandated to track and record technical data sheets and non-conformances in a standardised manner.

Because the data collected for analysis in this research was project specific, it cannot be used to analyse and generate conclusions that relate to different project circumstances. Data collection was also not the major scope of this research but was obtained from construction professionals at the Hwange Expansion Project site in the Matabeleland Province of Zimbabwe. The data was gathered from the various site personnel including site engineers, forepersons, safety, health, environmental, risk, and quality (SHERQ) practitioners as well as site operatives.

1.9 Research limitations

The major research limitation was the outbreak of the Covid 19 pandemic that resulted in the temporary closure of the project site and suspension of construction works delaying the data collection process. The pandemic also delayed the conduction of interviews with some having to be carried out on virtual platforms. Some of the interviews conducted online were interrupted by erratic internet connections.

The source of primary data was also limited to one construction project. This limited the identified rework risk triggers to one project locale and type. Furthermore, the expert knowledge derived from the interviewed professionals is not industry based but rather project specific. A diverse source of historical data would have aided in the development of a robust predictive model.

Due to the lack of sufficient historical project data, a fictitious dataset had to be compiled based on the survey results obtained from experts in concrete construction. The fictitious dataset comprised of randomly generated instances which, in some cases might not be a true reflection of situations on a real construction site. However, it was used for training the classification algorithms as well as developing the predictive model as it was the main focus of the study. The predictive model was theoretically evaluated and validated and could not be tested on site.

1.10 Key assumptions

It was assumed that the results of this research are a representation of the Zimbabwean construction industry. These results are not necessarily generalizable for other developing countries due to possible differences in the level of education and construction guidelines. It was assumed that all the respondents answered truthfully based on their personal professional opinion. The model is designed for use during the project construction phase (casting of reinforced concrete structural elements) and does not incorporate rework errors resulting from:

1) Changes and/or errors in design that affect activities during construction.
2) Project scope changes due to errors from the design or the contractor.
3) On-site fabrication errors that affect construction activities.

1.11 Key terms

1) **Algorithm**: A procedure that uses data to create a machine learning model which learns from the input data and recognises any patterns and trends the data might exhibit.
2) **Construction activities**: The processes involved in the construction of reinforced concrete structural elements.
3) **Construction errors**: Errors that occur during construction activities.
4) **Construction field rework**: The corrective work done during the construction of structural elements due to a non-conformance excluding the influence of design changes and errors as suggested by (Fayek, et al., 2003).
5) **Construction rework risk triggers**: The various factors, variables, and events that cause or initiate the occurrence of rework during the construction phase.
6) **Machine learning:** The study of algorithms that improve automatically through experience and by the use of data.
7) **Quality Control:** Control mechanisms and processes by which controls are carried.

1.12 Ethical considerations

To meet the requirements of internationally accepted ethical codes of practice and standards, an ethical application procedure was filed and accepted by the University of Stellenbosch with the reference number ING-2019-11592. The following measures were observed during the data collection process:

1. Names of research participants were not recorded on research instruments to maintain anonymity.
2. No form of monetary remuneration or otherwise was awarded to any respondent or participant in the research.

3. The research information obtained from the organization was kept confidential by not discussing or distributing information to any person other than the researchers' supervisor.

Standard quality assurance was also observed during the interviews in terms of the general conduct and interviewer competence, quality of data captured as well as accuracy in data analytics.

1.13 Research layout

This study is presented in seven chapters with the brief descriptions presented below:

Chapter 1: **Introduction** - the introductory chapter comprises the research background, problem statement, research motivation, justification of the research, the aim and research objectives, research design, research scope and limitations, key assumptions made in the study, the definition of key terms as well as the general thesis layout.

Chapter 2: **Literature review** - the literature study section discusses rework definitions, risk triggers that initiate rework occurrence, the negative impacts of construction field rework on project cost, schedule, and quality as well as the various rework migratory strategies developed over the years. The section also provides a brief overview of the different classification algorithms and how they function in prediction estimations.

Chapter 3: **Research methodology** - the research methodology provides information on the research design, data collection methods, tools, and accessories used in the research together with data analysis techniques adopted in the study.

Chapter 4: **Exploratory study** – this chapter presents the background and quality management systems implemented in the exploratory study which is the primary source of information for this research.

Chapter 5: **Data analysis and processing** – this chapter reports the analysis and interpretation of the data gathered represented in the form of tables and charts. The preparation and processing of the analysed data to develop a dataset used for modelling is also discussed.

Chapter 6: Modelling, evaluation, and deployment – this chapter reviews and evaluates the classification algorithms used in modelling. Additionally, the development of a predictive model that can be used by individual organizations in the management of structural concrete rework is discussed.

Chapter 7: Conclusions and recommendations – closing off the report**,** this chapter summarizes the major findings of the study together with recommendations for future research.

CHAPTER TWO

LITERATURE REVIEW

2.1 Chapter Overview

The review contained in this chapter aims to provide an appreciation of rework literature in the construction industry. To achieve the aim of the research, it is prudent for one to understand the common errors that lead to structural concrete rework on site. Rework covers quite an extensive scope, therefore for the purposes of this research, a working rework definition is established. Rework risk triggers specific to reinforced concrete element construction are identified and categorized into phases that will aid in the development of the predictive data analytics tool discussed later in this report. The negative impacts of these risk triggers are also discussed as the minimization and management of these form the framework of the study.

Discussed in section 2.6 are some of the rework management strategies developed by other researchers in the field. The last point of discussion proposes the application of supervised machine learning algorithms in rework management which seeks to bridge the gap in knowledge between rework management and predictive data analytics in the construction industry for structural concrete elements. This will promote data-based decision-making in the management of rework in the construction industry.

2.2 Introduction

The construction industry interchangeably uses the words “defect”’, “quality deviation”, “snag”,” rework”, “error” and “quality failure” to describe imperfections in reinforced concrete structure construction according to (Love et al., 2000; Josephson et al., 2002; Love, 2002; (Sommerville and McCosh, 2006). These terms are subjective and relate differently to various professionals but ultimately translate to client dissatisfaction with the project outcome (Llozor et al., 2004). If these terms are not defined clearly, inaccurate, ineffective, and inappropriate strategies will be developed to prevent and reduce their occurrence. Davis et al. (1989), suggests that non-conformances and defects are the same phenomena. However, ISO 9000 (2005) defines a non-conformance as “failure to adhere to stipulated standards” and defects as “non-fulfilment of a stipulated requirement relating to specific tasks”. It can be established that when a project together with its components are considered complete but differ from the set standards or requirements resulting in the decision to either accept or rectify, is grounds for a non-conformance. Battikha (2008), suggests that the term “defect” can be used whenever there is a failure in performance or client requirement of a building that develops structurally or through building services.

Whichever the case, the occurrence and corrective action taken to rectify a non-conformance or defect is regarded as “rework”. The causes of rework are a result of quality deviations according to Burati et al. (1992), Construction Industry Institute (1989) and Davis et al. (1989) non-conformance according to Abdul-Rahman (1995), defects according to Josephson and Hammarlund (1999), and failure in quality (Barber et al., 2000).

2.3 Rework definition

2.3.1 General

Construction literature over the years has provided several rework definitions and interpretations. It can be established that there is no standardized or consistent rework definition. According to Love (2002), some of the rework interpretations derived include quality deviations and failures, defects, as well as non-conformances. Love (2002) further ascertains two main definitions of the term which define rework as the process whereby an item is corrected so that it conforms to the required quality Ashford (1992) and the act of redoing an activity due to failure to conform to stipulated standards (Construction Industry Institute, 1989). Love et al. (1999a) maintained that rework is a result of customer dissatisfaction and is also associated with “waste” since the occurrence of rework-related activities does not add value to construction projects.

Most studies reveal that rework is greatly influenced by human error Love (2019) including and is not limited to faulty judgements, omissions, or lack of skill. Love and Li (2000) are of the opinion that rework can be defined as corrections of deficits. Love et al. (1999) put forward that rework is a result of organizations not being quality oriented. However, Love et al. (2000) refer to rework as the unnecessarily redoing of an activity that was implemented incorrectly in the first trial which usually incorporates design related errors and changes that are rectified during the construction phase. Contrastingly, Fayek et al. (2003) define rework as the total direct costs associated with the re-doing of construction work irrespective of the trigger excluding change orders and offsite manufacturing errors. Table 2.1 summarizes the various rework definitions derived from literature.

Table 2.1; Various rework definitions derived from literature.

Source	Rework Definition
Hammarland and Josephson, 1991	Quality failures.
Ashford,1992	Process whereby items are corrected so that they conform to the initial requirement.
Burati et al.,1992	Deviation in quality.
Abdul-Rahman, 1995	Non-conformance
Barber et al.,2000	Unneeded effort of repeating tasks that were incorrectly executed the first time.
Love and Li, 2000	Quality failures.
Alwi et al, 2001	Events that are non-value adding and negatively impact project performance.
Josephson et al., 2002	Unnecessary works done to correct construction errors.
Fayek et al., 2003	Field activities implemented more than once removing construction works that have already been implemented the first time.
Love and Edwards, 2004a	Non required effort of redoing activities that were wrongly implemented the first time.
Hwang et al.,2009	Works redone due to requirements not being met.
Zhang, 2009	Doing something more than once due to non-conformance.

2.3.2 Construction Field

The preceding definitions define rework in a broader perspective whilst this study seeks to understand the immediate causes of construction field rework. Field rework is defined as any construction activity/activities done more than once or activities that eliminate works already done in a project (Construction Industry Institute, 2001a). Rogge et al. (2001) narrow down rework to the field activities that are done more than once removing previous work as part of the project. In this research study, rework will be defined as the corrective work done during the construction of reinforced concrete structural elements due to a non-conformance excluding the influence of design changes and errors as suggested by (Fayek et al., 2003).

2.4 Rework risk triggers

An event or condition that initiates the occurrence of risk is commonly termed a risk trigger. Spacey (2016) defines a trigger as the root cause of a particular event. Based on this definition, rework risks triggers will be defined as the various factors, variables, or events that cause or initiate the occurrence of rework for the purposes of this research. Rework is predominantly caused by human errors which include rule or knowledge-based mistakes, lapses of attention, as well as acts of omission and commission (Love and Josephson, 2004 ; Love and Sing, 2011; Love and Li, 2000and Prodonovich et al., 2014). Similarly, Love, Skitmore and Earl (1998b) document that rework results from errors, omissions as well as changes during the project cycle.

Love et al. (2009) argue that the likelihood of rework occurrence increases the longer an error goes undetected which in turn significantly impacts project cost, schedule, and quality. These errors can be a result of an array of complex interactions hence an attempt to isolate one variable can significantly minimize rework occurrence according to Love et al. (2009). Construction field rework can also be triggered by a variety of factors/variables which include changes, defects, errors, failures, and other non-conformances or deviations in quality. Other triggers include constructability issues, unsuitable environmental and site conditions as well as omissions and or errors in the execution of construction activities. The following section will discuss the various causes of rework triggers in detail.

2.4.1 Rework risk triggers and classification

According to Koskela (1992), uncertainty is the major cause of construction rework. Similarly, Huovila and Koskela (1997) emphasised that this uncertainty is human in nature and springs from missing, unreliable, inaccurate, and conflicting information. Huovila and Koskela (1997) opine that uncertainty results from a variety of interrelated factors that are not necessarily poor information. Therefore, if rework is to be effectively prevented and reduced, it is imperative for one to identify and understand its root causes (Rodrigues and Bowers, 1996).

Rework can occur at any phase in the construction process. Love and Edwards (2004a) categorized rework root causes into three different factors which are

1) Client-related factors
2) Design-related factors
3) Contractor-related factors

They further developed a model based on these three factors. In their model, they integrated the various project and organizational characteristics that cause rework as shown in Figure 2.1.

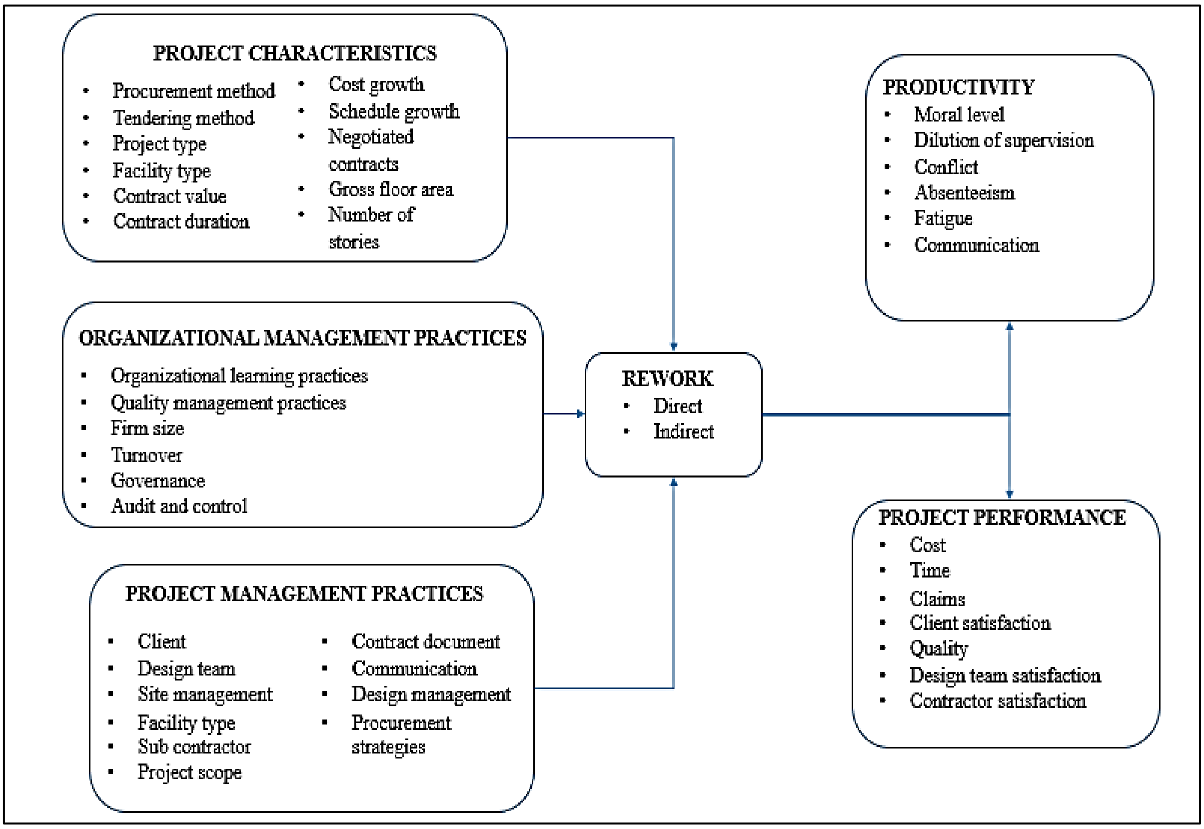

Figure 2.1; Project and organizational characteristics that cause rework (Love and Edwards, 2004a).

Similarly, Mastenbroek (2010) determined the rework causes in construction projects based in Honduras, Central America, and the most frequent rework causes were reported as shown in Table 2.2.

Table 2.2; Rework causes in the design and construction (Masterbroek,2010).

Project Phase	Rework Trigger
Design	Change by contractor
	Change by owner
	Finance related changes
	Economic related changes
	Lack of coordination
Construction	Changes by client
	Extra orders by client
	Plant breakdown/defects
	Late material delivery
	Change in construction methodology to enhance constructability
	Change in construction methodology due to unfavourable site conditions
	Noncompliance in construction quality

Hwang et al. (2009) affirmed that for effective management and elimination of rework, its risk triggers need to be fully understood and classified accordingly. Various literature sources discuss how rework is a result of the construction cycle's intricate characteristics. Similarly, Burati et al. (1992) developed a classification system comprising of three stages. The first stage or level defines the particular project phase affected i.e., pre-planning, design, construction, operational etc. The second stage determines the trigger of the rework (change or error) and the rework type. The last level classifies whether the resultant rework is under construction damage or change improvement (Nesan and Holt, 1999). Research conducted by Burati et al. (1992) categorized rework in five major groups as shown in Table 2.3 although it can also occur in management and administration.

Table 2.3; Rework classification (Burati, et al.,1992).

Phase	**Deviation Category**	**Description**
Design	Design Error	Error made during design
	Design Omission	Omission made during design
	Design Change/Construction	Design changes made at the request of the field or construction personnel
	Design Change/Field	Change due field conditions, a deviation could not have been foreseen by the designer
	Design Change/Owner	Change initiated by owner (Scope definition)
	Design Change/Process	Change in process, initiated by owner/designer
	Design Change/Fabrication	Change initiated or requested by fabricator or supplier.
	Design Change/Improvement	Revisions, modifications and improvements in design
	Design Change/Unknown	Redesign due to an error
Construction	Construction Change	Change in construction methodology to enhance constructability
	Construction Error	Results of erroneous construction methods
	Construction Omission	Omissions of some construction activity or task.
Fabrication	Fabrication Change	Change made during fabrication
	Fabrication Error	Error made during fabrication
	Fabrication Omission	Omission made during fabrication
Transportation	Transportation Change	Change made in method of transportation method
	Transportation Error	Error made in transportation method
	Transportation Omission	Omission made in transportation
Operation	Operability Change	Change made to enhance operability

Feng et al. (2008) reported that rework can either be viewed positively or negatively. Whenever designs are reworked during design reviews and the project team has a better understanding of the client's requirements, this positive form of rework adds value to the project. Love et al.(1997) conducted research on two construction projects and developed a system of classifying rework which is based on three fundamental groups which are people, design and construction. Love et al. (1997) came to the conclusion that some of the rework causes are almost always correlated due to the level of complexity of construction activities. A variety of rework causes are included in each category enhancing the understanding of the systems' user. Based on this classification, Alwi et al. (2001) developed a rework cause classification system shown in Figure 2.2. Alwi et al. (2001) conclude that "people related" rework accounts for approximately 60% of total rework costs.

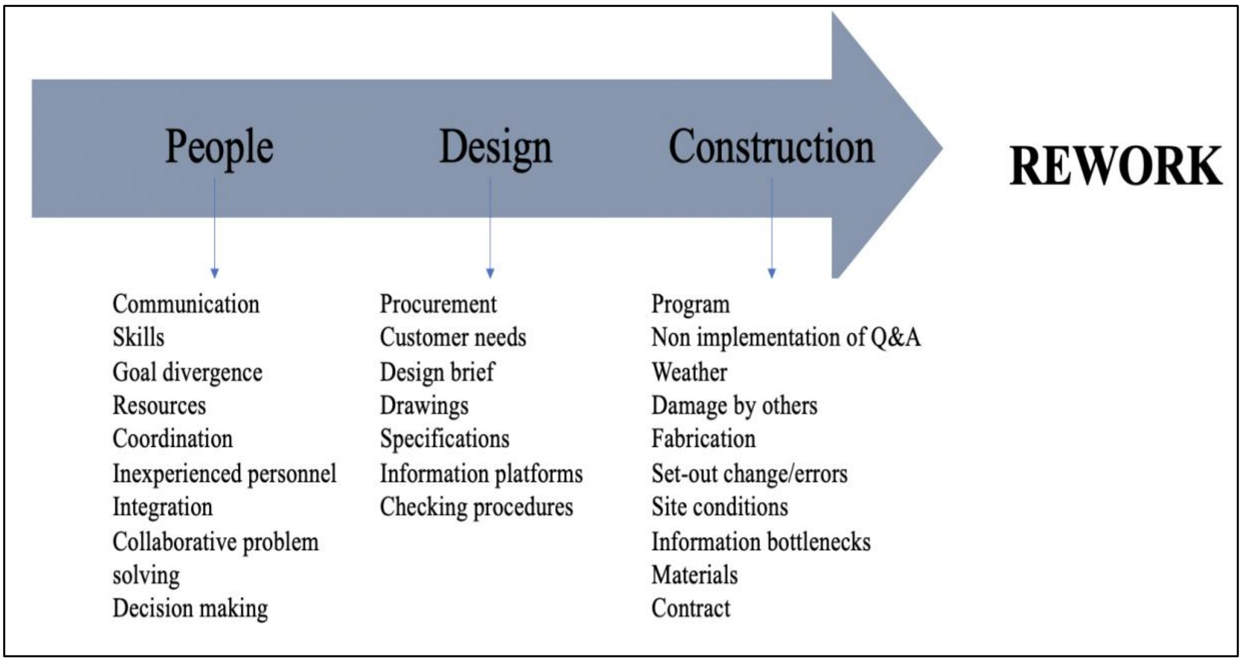

Figure 2.2; Rework cause classification (Alwi et al., 2001).

Similarly, Love and Li (2000) grouped rework causes into client-initiated changes, non-variations as well as defects. The Construction Owners Association of Alberta (2001) developed a more detailed "fish-bone" classification system. This classification system is sometimes referred to as the "cause and effect" diagram that explores the actual rework causes as shown in Figure 2.3. The fish-bone system is made up of five key rework areas and the possible causes associated with each key area.

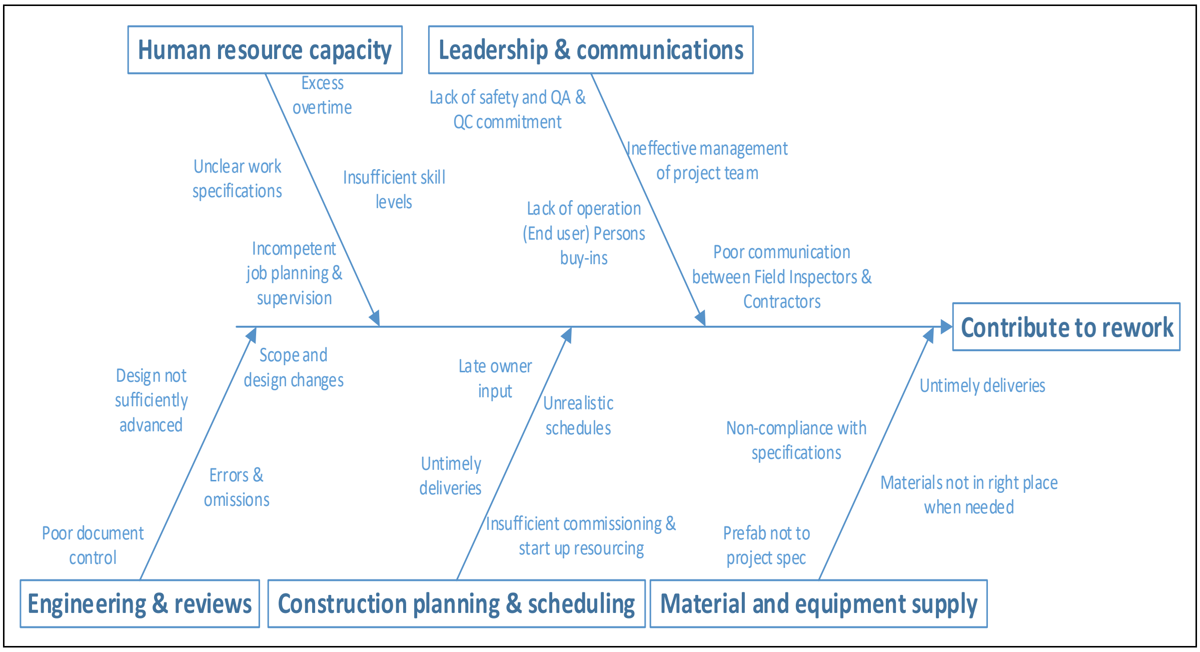

Figure 2.3; Rework fishbone classification (Fayek at el., 2003).

2.4.2 Construction rework risk triggers

The previous section broadly identified rework risk triggers for the reader to appreciate how literature categorizes them. This section specifically aims to achieve the first and second objectives of the study by identifying the various construction risk triggers that result in rework of structural concrete elements. Categorization of these risk triggers is based on research by Mastenbroek (2010) and will assist in the development of the predictive model. The study by Mastenbroek (2010) is based on the construction of three 8 storey residential buildings with field rework data having been collected over a period of 3 months, approximately a quarter of the construction period. Similar to this research, the study also investigated structural concrete elements such as beams, columns, and footings on an active construction project.

Addis (2005) is of the opinion that construction errors are a result of ineffective quality management systems. A South African based study conducted by Solomons (2014) revealed that approximately 80% of concrete corrosion is a result of construction errors such as poor workmanship. This study through the use of questionnaires and interviews identified the various rework risk triggers that significantly affect the quality of structural concrete beams, columns, and staircases in construction and concluded that the identified triggers are primarily caused by labour, management, and subcontractors.

Successful management of rework during construction can be accomplished by the adoption of effective quality management as mentioned in the previous sections. During the construction process, quality monitoring is a tool that ensures effective and improved quality control. A study conducted by Smallwood (2000) investigated the effectiveness of the various quality monitoring techniques that are typically adopted in construction projects. Table 2.4 summaries these quality monitoring techniques by presenting the effectiveness of the system by a rating system.

Table 2.4; Ranking of various quality monitoring techniques in construction (Smallwood, 2000).

Practise/System	Mean score					Rank
	Client	Designer	PM	Grade (5-9)	Mean	
Inspections/visual checks	4.21	4.67	4.17	4.43	4.37	1
Coordination meetings	3.93	4.44	4.00	4.29	4.17	2
Client briefing	3.93	4.25	3.33	4.05	3.89	3
Samples/references	3.21	4.00	3.50	4.43	3.79	4
Checklists	3.79	4.11	3.50	3.71	3.78	5
Close out reports	3.79	3.71	3.17	4.05	3.68	6
Tests	4.07	3.56	3.33	3.35	3.58	7
Value management	3.14	3.56	2.80	4.38	3.46	8
Constructability reviews	3.21	3.50	3.17	3.95	3.43	9
Document quality management system	3.31	3.22	3.00	4.00	3.38	10

Table 2.4 ranks inspections and visual checks as the most effective quality monitoring technique. Later in this study, data analytics will be investigated as a possible technique to improve concrete construction quality. Narrowing down to routine quality inspections, these checks are conducted throughout the concreting process.

Risk triggers can be categorized into phases during the construction of a reinforced concrete structural element. Categorization of these triggers simplifies the identification of typical construction errors that can occur and will also be adopted in the development of the machine learning model for this research. For this study, the risk triggers were categorized into three phases as shown in Figure 2.4.

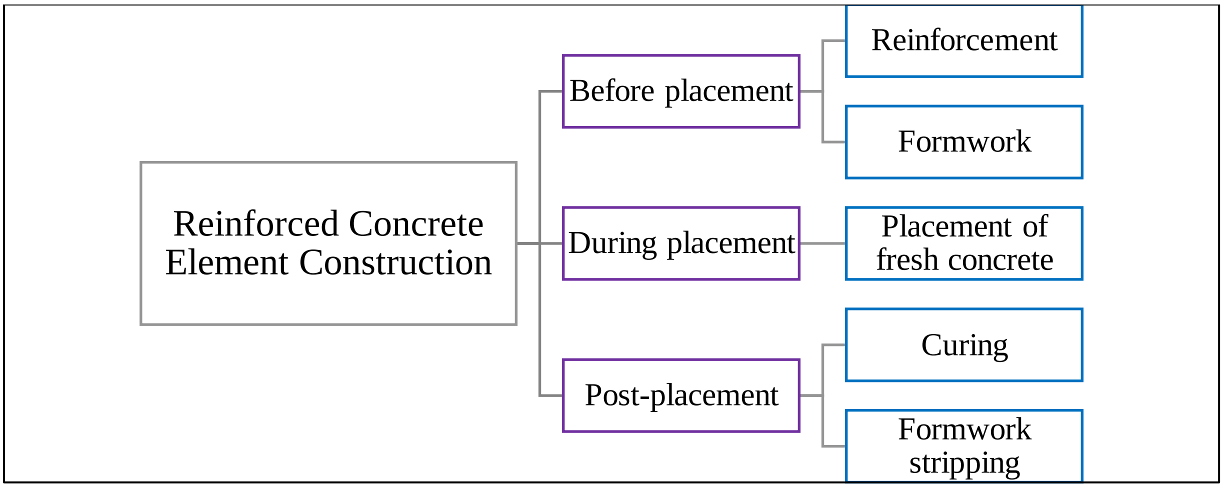

Figure 2.4; Phases involved in reinforced concrete element construction.

The first phase which is prior to the placement of concrete as presented in Figure 2.4 consists of the steel fixing and formwork placement processes. Steel fixing entails the process of placing and securing steel reinforcement bars in reinforced concrete construction. The following risk triggers that can potentially lead to rework during steel fixing have been identified (Addis, 2005):

1) Insufficient tensile strength
2) Poor surface quality of reinforcement bars (degree of rusting for example)
3) Incorrect rebar spacing
4) Incorrect rebar size
5) Incorrect cover to reinforcement
6) Incorrect overlap length
7) Loose ties
8) Incorrect link spacing
9) Incorrect splice positioning

Formwork placement is a pre-placement process that entails the erection and bracing of temporary or permanent moulds into which the concrete will be cast. Formwork can be fabricated from a variety of materials such as timber, steel as well as glass fibre reinforced plastics. During the process of formwork installation, the following construction errors can contribute to rework on site according to Manuel et al (2018):

- Incorrect final elevation
- Insecurely fixed/secured formwork
- Axis displacement

- Incorrect formwork sizes
- Poor formwork surface quality
- Damaged and deformed formwork boards
- Unacceptable formwork tolerances.

When formwork is fixed and secured in place, concrete placement can be effected. Ready mix concrete trucks and concrete pumps are usually used during the casting of concrete in particular moulds. Neville (2010) establishes that in the absence of concrete retarding agents, the mix must be placed within a fifty-minute window after batching to avoid the formation of cold joints. Furthermore, absorbent surfaces in contact with the mix must be adequately wetted before concrete placement. During the concrete placement process, the vibration of the concrete using various methods is done to ensure the dissipation of air voids that weaken cast concrete.

During the concrete placement process, the following risk triggers were identified according to Neville (2010):

- Incorrect mix design
- Failed slump
- Inadequate vibration
- Over vibration
- Placement during adverse weather.

The post placement process includes activities such as curing, protection, as well as formwork stripping of the cast concrete after the element has gained sufficient strength. Curing is achieved by keeping the surface moist/ preventing the loss of moisture through the use of a variety of protective barriers. The following discrepancies may lead to potential rework after concrete placement is complete according to Odgers et al. (2012):

- Incorrect final elevation levels
- Inadequate curing
- Poor surface finish
- Damage during stripping

2.4.3 Effects of construction risk triggers

It has been established from the previous section that poor construction quality and construction errors lead to structural concrete rework. A number of projects have failed as a result of poor construction quality practices. A UK based study by Shammas-Toma et al. (1996) revealed that a number of concrete bridges required rework due to corrosion resulting in millions of pounds being spent in remedial works. The study recommended the adoption of enhanced construction quality which results in improved quality of workmanship as well as a reduction in rework occurrence. Similarly, a shopping centre in Pretoria North recorded a case of scaffolding collapse due to poor quality management according to (Smallwood, 2012).

The risk triggers identified in Section 2.4.2 potentially lead to various structural concrete rework during its fresh or hardened state. Adopting the proposed reinforced concrete construction phases mentioned in the previous section, this section will discuss the effects of the previously identified risk triggers. Solomons (2014) conducted a study on structural concrete beams, columns, and staircases and established that the quality of structural concrete works is greatly influenced by cover to reinforcement, steel spacing as well as kicking of formwork due to inadequate bracing.

Helmy (2017) affirms that during steel fixing, inadequately secured or misplaced steel reinforcement bars usually result in structural cracking because the steel will not be functioning as structurally intended. Sometimes, if steel reinforcement bars are not securely tied, the bars tend to migrate towards the surface, compromising the concrete cover which usually results in corrosion of the steel bars. Congestion of reinforcement bars can also occur if they are inadequately secured and tied making it difficult to effectively compact the concrete even though it may be workable.

A field study conducted by Smith (2010) recorded the number of non-compliances to design specifications in beam and slab measurements during inspections. The study concluded that approximately 44% and 23% of the slab and beam measurements respectively did not comply with SANS 2011-CC1 (2007). Ronne (2006) also investigated variation in concrete cover to reinforcement in the South Africa industry and concluded that 30% of the concrete cover data did not conform to the stipulated requirements. Sagging of reinforcement bars can also occur due to loosely tied rebars as shown in Figure 2.5, leading to reductions in effective depth of sections compromising the durability of the entire element.

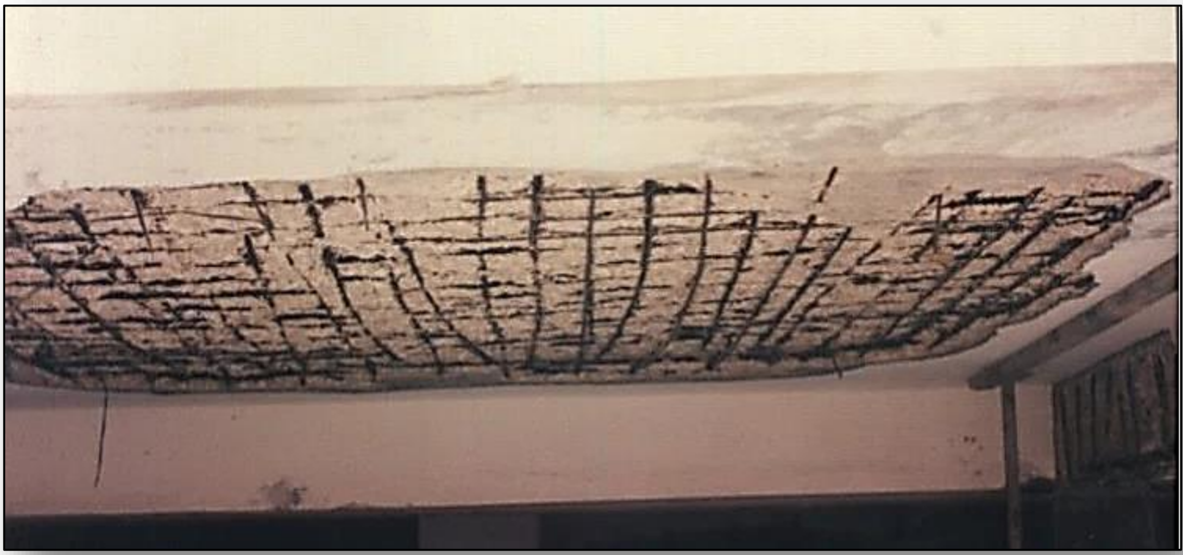

Figure 2.5; Sagging reinforcement bars reducing the effective depth of structural elements (Hemly, 2017).

During the installation of formwork, if it is improperly aligned, this might result in irregularities of the concrete surface after it is cast. If this occurs to members in contact with a flow of water, cavitation erosion may occur. If the formwork is inadequately sealed, this might induce grout loss during placement compromising its overall strength. Movement of the formwork during vibration may result in formwork collapse. If the formwork is removed prematurely (before sufficient curing) the concrete because it has not gained enough strength and may develop cracks due to overstressing. A study by Manuel et al. (2018) established that formwork can contribute up to 36% of the collapse of structural concrete elements during construction by using old shutter boards, using shutters that are not properly cleaned as well as failure in the shoring system as shown in Figure 2.6.

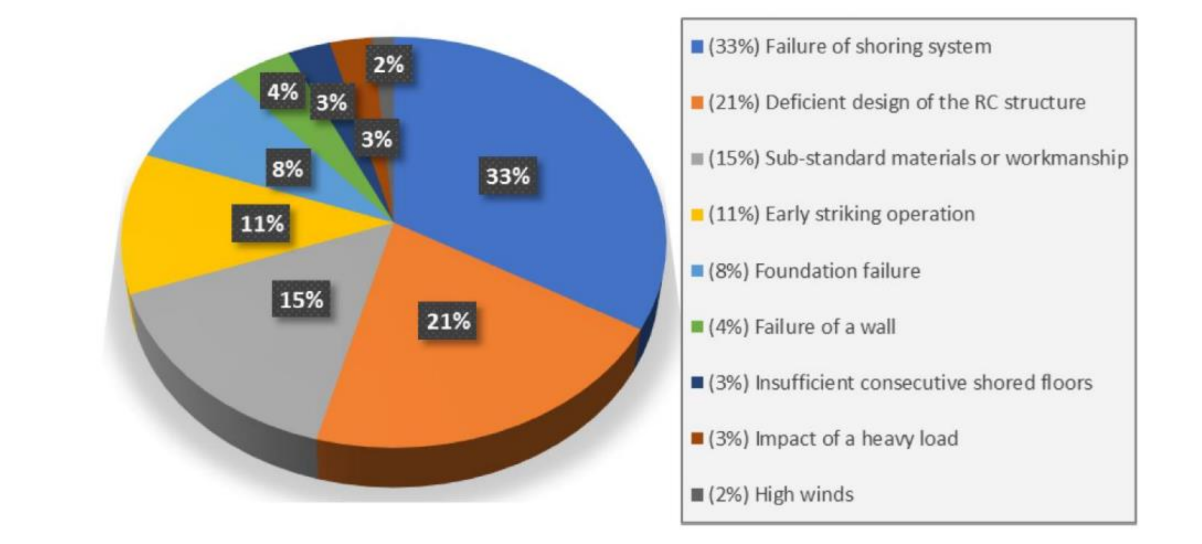

Figure 2.6;Factors that contribute to structural failure during construction (Manuel et al., 2018).

After engaging in a brief discussion with the senior civil engineer at a construction site in Hwange, the engineer was of the opinion that during concreting, water might be added to the concrete mix onsite to increase workability but this might result in reduced concrete strength and durability because the water to cement ratio is inversely proportional to concrete strength and durability. Neville (2010) also established that some concrete defects that lead to rework are a result of incorrect concrete casting or the failure to remove air effectively during compaction. During a site visit, Solomons (2014) concluded that honeycombing in structural elements is a result of inadequate compaction during placement. The same study also established that grout loss experienced in structural concrete columns was a result of inadequately sealed formwork. If concrete is improperly vibrated, this might result in honeycombing as shown in Figure 2.7. Similarly, Addis (2005) states that in order to prevent honeycombing, concrete should be sufficiently poked. Over compaction on the other hand can lead to settlement whereby coarse aggregates settle allowing the concrete paste to rise to the surface.

The durability of cast concrete is also compromised if it is not adequately cured at the required temperature and humidity. Hairline cracks and surface disintegration are some of the surface defects associated with improper curing (Helmy, 2017). Additionally, the concrete cover is affected by poor curing leading to poor durability of concrete according to Addis (2005).

Figure 2.7; Honeycombing (indicated by the red arrow) as a result of insufficient poking during placement of fresh concrete (Helmy, 2017).

To achieve the aim of this study, Table 2.5 summarizes the consequences of the identified risk triggers in section 2.4.2 on reinforced concrete structural elements to better understand how to effectively manage rework.

Table 2.5; Summary of effects of rework risk triggers on concrete elements.

Phase	**Rework Risk Trigger**	**Effect on element**
Pre-placement	Incorrect placement of steel reinforcement.	Reduces concrete cover exposing rebars to environmental conditions increasing chances of corrosion.
		Overall strength of structural member is also compromised.
	Insufficient concrete cover	Exposes rebars to corrosion resulting in deterioration and spalling of concrete.

	Rework Risk Trigger	**Effect on element**
Pre-placement	Inadequately sealed formwork joints	Grout leakage occurs resulting in a decrease of overall concrete strength and a highly porous element.
During placement	Inadequate compaction	Creates bug holes and honeycombs that make the element porous.
	Over compaction	Settling of coarse aggregates resulting in decrease in overall strength and bleeding.
	High water to cement ratio	Increases permeability of element and reduces strength.
Post-placement	Improper curing	Produces surface cracks as well as reduces the strength of the member.
	Mechanical damage during formwork stripping	Irregular surface appearance
	Premature stripping of formwork	Produces surface cracks as well as reduces the strength of the member.

2.5 Impacts of construction field rework

Rework significantly contributes to time/ schedule wastages and overruns according to Chan and Kumaraswamy (1997) and the Construction Industry Institute (2001b) which inevitably negatively impacts the final project costs, resources, and quality (Love and Edwards, 2004a). Construction field rework may also trigger monetary claims for the time spent rectifying or redoing in the event that the contractor, for example, seeks some form of remuneration from the responsible parties (Palaneeswaran, 2006). Cooper (1993) reported that rework emerges in the form of overtime which results from the hiring of extra resources (labour and plant), compromised project quality, and reduction in scope. Lower profits, drops in market shares, tainted organizational reputation, as well as reductions in productivity are among some of the adverse effects of rework according to (Cooper, 1980; Ackermann et al., 1997; Eden et al., 2000).

2.5.1 Cost

Due to high levels of subjectivity in rework definitions and interpretations by construction experts and researchers, standard rework data is usually difficult to obtain resulting in inconsistent rework data collection and quantification (Love and Smith, 2003). Josephson and Hammarlund (1999) suggested rework costs in building construction projects be it residential, industrial, or commercial ranged from 2% to 6% of the contract values, while Fayek et al. (2003) reported that this figure ranged from 2% to 12%. Similarly, Love and Li (2000) established that direct costs of rework in residential projects can be estimated to be 3.15% of the contract value and 2.4% in industrial building projects. Furthermore, Oyewobi and Ogunsemi (2010) established that for new buildings, rework attributed 5.06% of the contract value and for refurbishment projects, an estimated 3.23% project cost overrun is because of rework. (Hwang, Zhao and Yang, 2018) established that rework can contribute up to 25% of the total project value.

However, in civil infrastructure projects, Love et al. (2010) reported that the cost associated with rework was found to be 10.29% and 16.5% of the contract value with the differences owing to the methodology adopted in capturing and analysing rework costs, which also influence determinations of the causal nature of the rework costs (Fayek et al., 2003; Love and Edwards, 2004a; Love and Sing, 2011). Mahamid (2016) conducted a study in West Bank, Palestine and concluded that for residential building projects, rework associated costs range between 10%-15% of the original value. In another research project based in Hong Kong, China, a private building with construction works worth $60 million resulted in 16.1% and 4.8% direct and indirect rework costs respectively. Various researchers have concluded that rework associated costs in poorly managed construction projects can be up to as much as 25% of the original contract value and up to 10% of the total project costs (Barber et al., 2000) and (Love and Li, 2000). The Construction Task Force in the United Kingdom according to Egan (1998) indicated that rework contributes up to 30% of construction costs and the United States of America based Construction Industry Institute estimated a US$15 billion annual loss attributed to construction rework for industrial construction projects (Construction Industry Institute, 2001a).

Burati et al. (1992) and Barber et al. (2000) indicate that in major civil engineering projects, rework attributes roughly 5 to 20% of the original contract value. Having suggested that rework costs can contribute up to 12.6% of project costs Love (2002a) later argued that some construction projects may result in higher rework costs due to the unique nature of every project. Marosszeky (2006) conducted a research in Australia and reported that a mean of 5.5% of the original contract value attributed to rework costs with a breakdown of 2.75% direct costs and 1% indirect costs. Low and Yeo (1998) indicated that a construction contractor typically runs 5-10% over budget due to rectifying poor quality construction works. Meshksar (2012) reported that rework can cost up to 1.30% - 3.30% of contract value with a mean of 2.095%. Oyewobi et al. (2011) investigated on an elemental basis and concluded that rework costs can be an average of 5.29% of the original contract value. Davis et al. (1989) however reported that additional costs of up to 12.4% of the total project cost are attributed to rework and Gunawardena et al. (2004) revealed that the cost of rectifying construction works that were not initially done up to standard can cost an average of 10% of the total contract sum.

A case study research based on 42 contractors revealed that costs of construction rework ranged from 2% to 5% of the contract value as reported by (Love and Li, 2000) and (Kakitahi et al., 2013). Considering indirect rework costs according to Barber et al. (2000) increased the contract value to a range of 16% to 23% which is inclusive of all delay costs incurred. However, if the delay costs are eliminated, the rework costs would contribute about 3.6% to 6.6% of the contract value. Research in Nigeria revealed that rework in a construction project caused a cost overrun of N3,341,805.00 (three million, three hundred and forty-one thousand, eight hundred and five Naira only) which was about 12.85% of the contract value. Another researcher reported construction rework costs stood to be around 2.75% of the original contract value, which is similar to the estimates suggested by Fayek et al. (2003), Love and Li (2000) and Josephson et al. (2002) but lower than that estimates by Love (2002), who used a questionnaire survey for data collection.

A study conducted by Rhodes and Smallwood (2003) revealed that rework costs can account for up to 13% of the value of completed construction. Barber et al. (2000) also acknowledged that the total costs of rework could be as high as 23% of the initial contract value. Marosszeky (2006) conducted research based in Australia and discovered that the mean of rework costs was found to be 5.5% of contract value including 2.75% as direct costs, 1.75% indirect costs for main contractors and 1% indirect costs for subcontractors. Wasfy (2010) conducted research on a residential commercial tower in Saudi Arabia and came to the conclusion that rework has the potential to increase the cost of the different work categories between 2% to 30%. It must be taken into account that due to the different definitions and interpretation of rework, differing project scope, and data collection methodology adopted, these numbers are not fully comparable. They do however give an indication of the extent to which rework influences cost overruns (Love and Sohal, 2013). A summary of the rework costs extracted from various literature sources is provided in Table 2.6.

Table 2.6; Summary of rework costs derived from literature.

Source	**Origin**	**Rework Cost**
BRE, 1981and BRE, 1982	UK	15% of total project cost
Hammarlund and Josephson, 1991	Sweden	4% of total project cost
Cnuddle, 1991	Sweden	10-20% of total project cost
Burati et al., 1992	USA	12.4% of total project cost
Abdul-Rahman, 1995	UK	5% of tender value
Love et al., 1998b	Australia	2.4-3.15% of contract value
Alwi et al., 2001	Indonesia	2.01-3.21 of total project cost
Barber et al., 2000	UK	16-26% of total project cost
Love and Li, 2000	Australia	3.15-12.4% of contract value
Love and 2002	Australia	6.4-5.6% of original contract value
Josephson et al., 2002	Sweden	2.3-9.3% of contract value
Rhodes and Smallwood, 2003	South Africa	13% of total completed project cost
Love and Edwards, 2004a	Australia	6.5% of contract value
Marosszeky, 2006	New South Wales	5.5% of contract value
Palaneeswaran, 2006	Hong Kong	5.2% of contract value
Zhang, 2009	Canada	0.93-1.35% of total project cost

Source	**Origin**	**Rework Cost**
Oyewobi et al., 2011	Nigeria	4.49% of total project cost
Meshksar, 2012	Iran	1.85-2.1% of total project cost
Podges, 2017	South Africa	4-6% of the total value of concrete works

2.5.2 Schedule

Apart from the cost implications rework poses on construction projects as previously discussed, rework can significantly contribute to schedule growth in construction projects. A study conducted by Love (2002) on 161 building projects in Australia revealed that an average schedule growth of 20% was attributed to rework. A survey conducted by the Navigant Construction Forum in 2011 with over 40 experienced construction professionals concluded that rework can contribute to about 19% of schedule delays if not effectively managed. Fayek et al. (2003) investigated schedule delays associated with rework by monitoring rework related activities starting from its initial identification to when the activity had been rectified to conform to its desired standard. Rework duration according to their research included the standby or relocation time, the time required to rectify the activity, and the gear up to carry on with the original scope of the activity as shown in Figure 2.8.

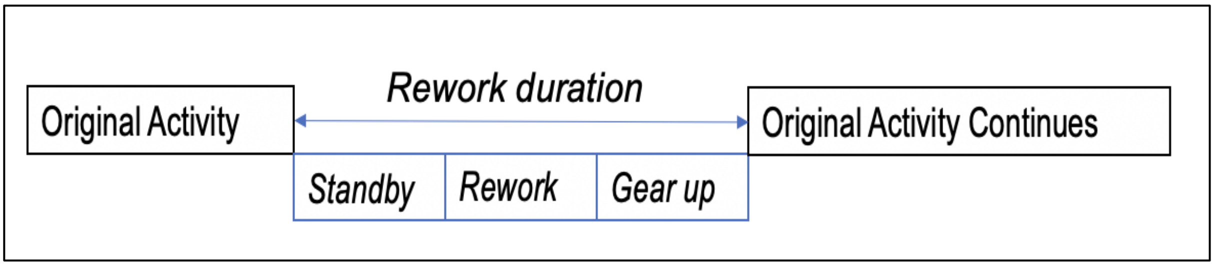

Figure 2.8; Time related components of rework (Fayek et al.,2003).

Abeku et al. (2016) reported a combined effect of rework on the projects' time and cost of 55% and 28% respectively. Another research indicated that, the time delay of rework from a case study and survey conducted was 4% and 5% of construction duration respectively (Meshksar, 2012). A Nigerian based research project on two building projects revealed that rework attributed a total of 43 extra days which translated to 38% schedule overrun in the first project and an extra 14 days which was 17% schedule overrun in the second project. This resulted in a mean schedule overrun of 15%, which is slightly lower than that reported by Love (2002) which was an estimate of about 21% of the contract schedule. A private building project in Hong Kong reported a time overrun of 277 days due to rework in which the original schedule at the award of the contract was 480 days which is higher than the 3- 8% range reported by (Meshksar, 2012). Research conducted by Wasfy (2010) on a residential-commercial tower in Saudi Arabia recorded an increase of project durations with a range of 10% to 77% resulting from project delays in the different rework categories.

2.5.3 Product Quality

Quality costs in construction have been captured by various systems developed over the years. .Cnuddle (1991) conducted a research on the costs associated with non-conformances on site and concluded that the cost ranges between 10% and 20% of the total project cost with 22% of the costs resulting from construction deviations. Similarly, the Construction Industry Institute (2005) concluded that on average, projects that fail to adopt a formal quality management system have rework costs that account for 6.5% of the contract value. According to Barber et al. (2000) and Love and Li (2000) rework costs in projects that do not implement effective quality management systems can be as high as 25% of contract value and 10% of the total project costs.

In addition, Love and Li (2000) affirmed that if a contractor implements a quality management system together with effective continuous improvement strategies, rework costs can be less than 1% of the contract value. According to Cusack (1992), projects that do not implement a quality system typically experience a 10% cost increase due to rework. The Building Research Establisment (1982) insisted on the utilization of effective quality control systems to save about 15% of construction costs by the elimination of rework. Kazaz et al. (2005) conducted research on a Turkish housing project that comprised 3100 residential units with up to five storeys. On average, the costs of quality failures for the project were found to be 11.5% of the total project costs. A summary of quality failure costs from various researchers is summarized in Table 2.7.

Table 2.7; Summary of quality failure costs from literature.

Source	Origin	Quality Failure Costs
Burati et al., 1992	USA	12.4% of the total project cost.
Abdul-Rahman, 1995	UK	5% of the project tender value.
Abdul-Rahman et al., 1996	UK	65% of the estimated project cost.
Hammarlund and Josephson, 1999	Sweden	4.9% of the production cost.
Barber et al., 2000	UK	6.6% of the budgeted project cost.
Love and Li, 2000	Australia	2.4% of the contract value.
Josephson et al., 2002	Sweden	2 % - 9% of the contract value.
Love and Edwards, 2005	Australia	12% of the contract value.

2.6 Proposed mitigation strategies

According to Yates and Lockley (2002) construction errors have become an inevitable reality in construction projects due to the innate human nature, however, the implementation of procedural inspections and effective communication during construction project execution can minimize such errors. Total quality management (TQM) processes that treat rework as a quality related issue are some of the conventional measures that have been proposed to minimize its occurrence. According to Davis et al. (1989), the lack of a robust quality management system may result in the inaccurate identification of quality deviations. Subsequently, the loss of information occurs, and activities/tasks that require improvement to minimize rework occurrence cannot be ascertained. Similarly, the Building Research Establishement (1982) demonstrated through various case studies the significant cost benefits of implementing a quality management system.

Quality costs as suggested by Campanella and Corcoran (1983) and as shown in Figure 2.9 are important although the reduction in quality failure costs will only reflect over time due to the time delay between cause and effect. Prevention and appraisal costs are unavoidable and should be borne by any organization aiming at delivering products or services right the first time. The elimination of rework causes significantly reduced appraisal costs as documented by (Low and Yeo, 1998). They established that a 1% investment in prevention costs can result in about 80% savings on failure costs. Therefore, in an attempt to minimize rework occurrences in the field, various approaches have been promulgated, some of which include visualization enabled technologies, lean construction as well as constructability reviews between the designers and construction site engineers.

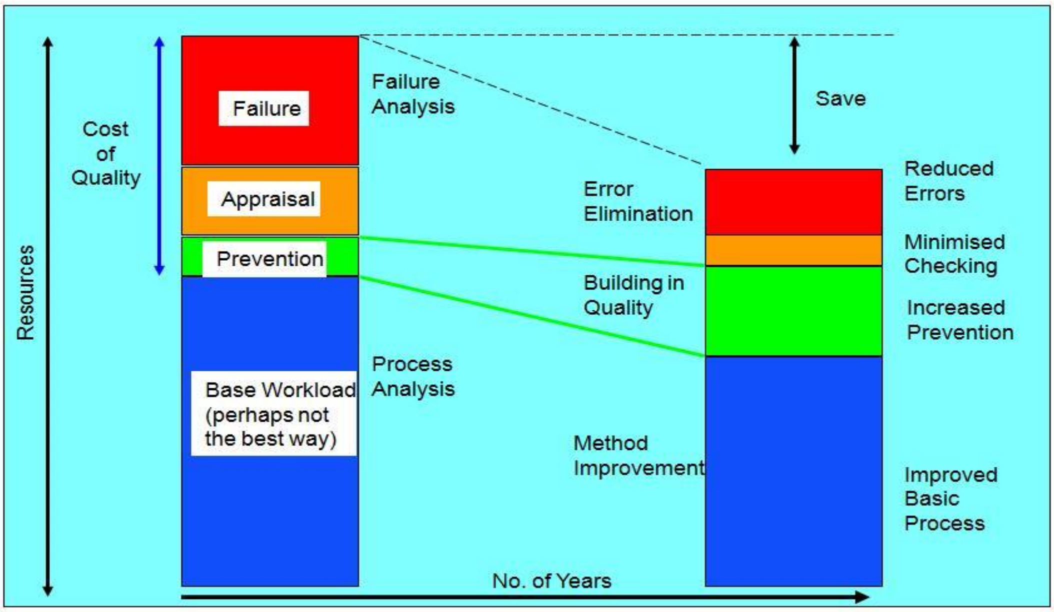

Figure 2.9; Benefits of investments in quality costs over time (Campanella and Corcoran (1983).

2.6.1 The Field Rework Index (FRI)

The Construction Industry Institute of Research developed The Field Rework Index (FRI) tool in order to proactively manage field rework (during preconstruction) in terms of early warnings as well as cost growth (Construction Industry Institute, 2001). Fourteen variables that represent various project characteristics were developed and tested with historical data from completed construction projects. The index however does not accurately predict the probability of field rework occurrence but provides early warning for any possibility of cost growth. To use the index, the user rates on a scale of 1-5, the 14 variables and sums the scores to get a unique FRI score. The FRI score is then compared to the various early warning levels displayed on the danger chart as shown in Figure 2.10.

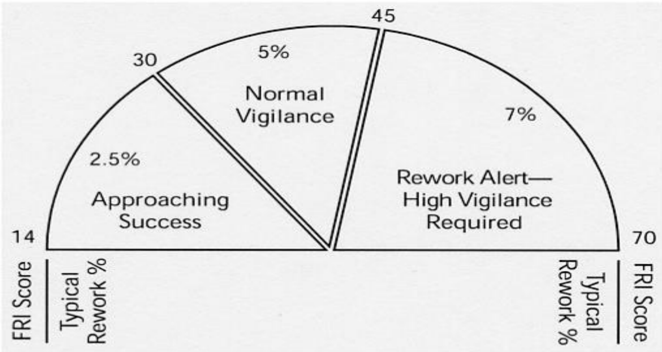

Figure 2.10; FRI Rework danger chart (CII;2001).

2.6.2 Zero field rework self-assessment opportunity checklist

The Zero field rework self-assessment opportunity checklist is a tool established by the CII Research Team to aid in the identification of any areas of improvement to further enhance construction site quality processes to achieve zero field rework (Construction Industry Institute, 2005). The tool developers are of the notion that previous construction site inspection checklists mainly focus on manuals/procedures, document and data control, as well as material control. The Zero Field Rework Self-Assessment Opportunity Checklist differs from other checklists in that it primarily focuses on human performance and takes a behaviour-based approach in rework reduction (Construction Industry Institute, 2005). The checklist comprises 8 elements and 117 questions with some of the elements being leadership, teamwork, employee involvement as well as communication. After the assessment is complete, project management then determines the most effective way to share the findings with the other team members and communicate the major learning point/s, implementation plan, and worker involvement (Construction Industry Institute, 2005).

2.6.3 COAA's Project Rework Reduction Tool (PRRT)

The Project Rework Reduction Tool "PRRT" is a commercial software established by the Construction Owners Association of Alberta "COAA" to measure project performance against predetermined rework causes. It is suggested that rework during the construction process can be prevented if an early warning system exists that identifies rework risk triggers (Construction Owners Association of Alberta, 2006). The tool is therefore designed to evaluate field rework and peg it against the major field rework risk triggers. The evaluation can be implemented at any point during the project cycle. The ratings are interpreted within the 5 principal rework cause classification of the COAA fishbone as shown in Figure 2.11. The mean rating may be used for calibration of projects of similar nature while the suggestion for practical solutions is given for continual improvement (Construction Owners Association of Alberta, 2006).

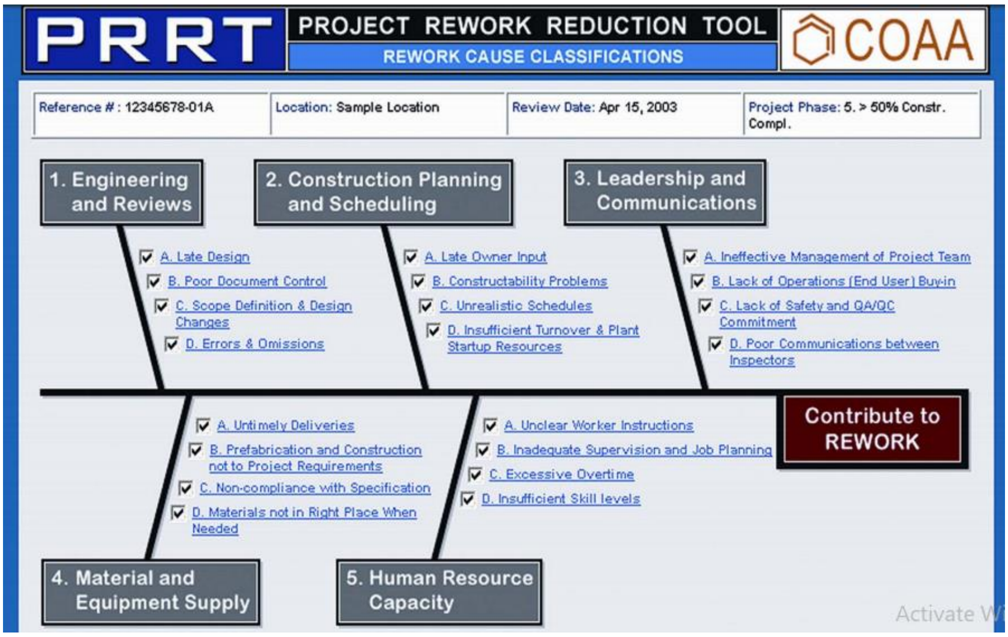

Figure 2.11; PRRT Program user interface (COAA,2006).

2.6.4 Ontario Pickering Nuclear Power Station Groups Rework Program

Zhang (2009) formulated a conceptual model that generalized rework reduction. The purpose of the model is to minimize and manage rework occurrence via te management of a continuous improvement cycle consisting of four key functional processes shown in Figure 2.12.

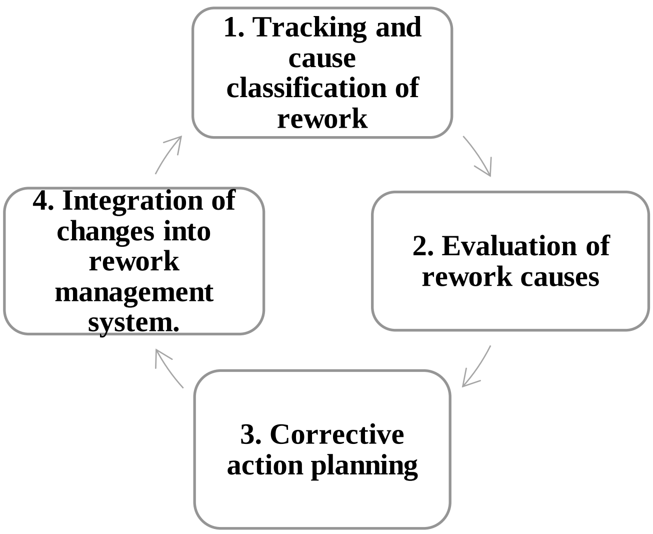

Figure 2.12; Continuous improvement cycle for rework reduction (Zhang, 2009).

Zhang (2009) further clarified that a critical factor in the rework reduction tool developed is the appropriate classification of rework risk triggers. This is due to the fact if rework risk triggers are defined and categorized in a formal structure comprising of practicable root causes, an increase in rework awareness is initiated. This generalized model defines and divides seven risk triggers into two groups as shown in Table 2.8.

Table 2.8; List of identified rework root causes (Zhang, 2009).

	Rework Risk Trigger	Description
		Process Group
1	Design and Engineering	• Errors and omissions in drawings and specifications. • Document control deficiencies. • Changes in scope. • Poor detailing.
2	Instruction and Inspection	• Ineffective communication. • Ineffective decision-making process. • Inadequate monitoring and control.
3	Schedule	• Field condition forecasting deficiencies. • Poor construction resource scheduling. • Lack of realistic work procedures.
4	Supply of material and equipment	• Misplacement and untimely deliveries. • Prefabrication defects. • Lack of advancements in tools and equipment.
		Human Performance Group
5	Knowledge	• Lack of sufficient knowledge to successfully complete task.
6	Skill	• Lack of domain specific skill. • Personal training deficiencies.
7	Self-discipline	• Non-conformance to rules, policies, work instructions or procedures. • Lack of motivation

Validation and evaluation of the rework program resulted in a decrease in the number of rework incidences reported by Ontario Nuclear. Rework incidences reported decreased from 1590 to 926 in the years 2006 to 2008. A decline in the percentage of rework labour hours from 1.06 % to 0.80 % for the same time interval was also observed. Correspondingly, for the same time interval, a decline in dollar value from 1.35% to 0.93% was reported. The above results verify the effectiveness of the rework reduction tool. A decline of 36% from 43% in human performance was documented for the same time frame which further validates how effective the rework reduction tool is, however, the researcher stressed the importance of improving human performance. Improving human performance has a positive effect on both safety performance and rework control.

2.6.5 Inter-project learning to avoiding rework

Mastenbroek (2010) analysed the various rework causes in five construction projects to assess the costs associated with rework in Honduras and suggested improvements that can be implemented to minimize rework costs. The rework causes identified are summarized into six categories which are:

1) Change orders
2) Coordination
3) Material deliveries
4) Construction methods
5) Personnel
6) Machinery

Mastenbroek (2010) further recommended that historical data collected (lessons learnt) from previous projects be implemented in new projects to improve productivity through the reduction of rework. Corrective actions compiled from one project can be implemented in future projects in the event that similar issues arise, provided the effectiveness of the corrective actions is verified. The adoption of inter-project learning promotes the identification of strategies to progressively generate, share and impact new knowledge (Mastenbroek, 2010).

2.6.6 Rework reduction through the use of construction lean improvement program (CLIP)

Basbeth (2014) conducted a study centred on the use of lean construction to reduce rework as summarised in Table 2.9. The rework reduction in this study is underpinned by the principle of lean thinking.

Table 2.9; Lean construction principles that aid in rework reduction (Basbeth, 2014).

Lean Principles	Description	Benefits
Lean Thinking	Based on the concept of lean production, lean thinking focuses of the achievement of the following three outcomes: • Maximizing value from the customers' view. • Waste elimination. • Creating smooth and reliable activity flow.	<u>Direct Benefits</u> • 16-40% increase in productivity • 25-48% decrease in lead time <u>Indirect Benefits</u> • Operational and organizational improvements. • Waste reduction. • Reduced delays and snags. • Improved team dynamics and skills transfer.

The research was based on a case study in Indonesia of a new Australian embassy building construction project. The frequency of non-conformance reports validated the effectiveness of adopting lean construction principles in rework reduction. The training of skilled labour to LCI standards, conducting of internal quality audits, and increased management attention were implemented for a period of 6 months which resulted in a 57% reduction of non-conformance reports and a project saving of 20 000,00 United States Dollars.

The tools and techniques discussed in this section have been developed to proactively manage rework occurrence in the construction industry. Some other methods summarized in Table 2.10 have also been developed to minimize the negative effects rework imposes on project performance and productivity.

Table 2.10; Summary of rework reduction tools and measures that have been developed over the years.

	Source	Proposed Rework Reduction Measure
1	Construction Industry Institute, 2005 and Construction Owners Association of Alberta, 2006	Effective management and leadership
2	Construction Industry Institute, 2005 and Construction Owners Association of Alberta, 2006	Sufficient and competent human resources
3	Construction Industry Institute, 2005	Involvement of all project members
4	Construction Industry Institute, 2005 and Construction Owners Association of Alberta, 2006	Effective communication
5	Construction Industry Institute, 2005	Effective teamwork
6	Construction Industry Institute, 2005	Proper project documentation
7	Construction Industry Institute, 2005 and Construction Owners Association of Alberta, 2006	Rework- Quality auditing
8	Construction Owners Association of Alberta, 2006 and Zhang, 2009	Competent supervisors
9	Construction Owners Association of Alberta, 2006 and Zhang, 2009	Continuous work evaluation throughout project cycle
10	Construction Owners Association of Alberta, 2006, Zhang, 2009 and Mastenbroek, 2010	Adherence to strict safety laws
11	Construction Owners Association of Alberta, 2006, Zhang, 2009 and Mastenbroek, 2010	Qualified and competent contractors
12	Construction Owners Association of Alberta, 2006, Zhang, 2009 and Mastenbroek, 2010	Project owner/client involvement
13	Construction Owners Association of Alberta, 2006 and Zhang, 2009 and Mastenbroek, 2010	Effective scheduling and planning
14	Basbeth, 2014	Using lean construction improvement

2.7 Machine Learning Overview

Machine Learning (ML) is the application of artificial intelligence that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Algorithms used for machine learning use computational methods, that allow ML models to learn from provided data instead of a predetermined governing equation. The ML algorithms are developed in such a way that they are adaptive in nature, and their performance improves with an increase in the number of data samples (MathWorks, 2020). For most machine learning problems, two techniques are usually used as shown in Figure 2.13. These are supervised learning which is based on training a model from a sample of known inputs and corresponding outputs. Unsupervised learning on the other hand tends to determine patterns hidden in data or intrinsic elements in the input data. (MathWorks, 2020)

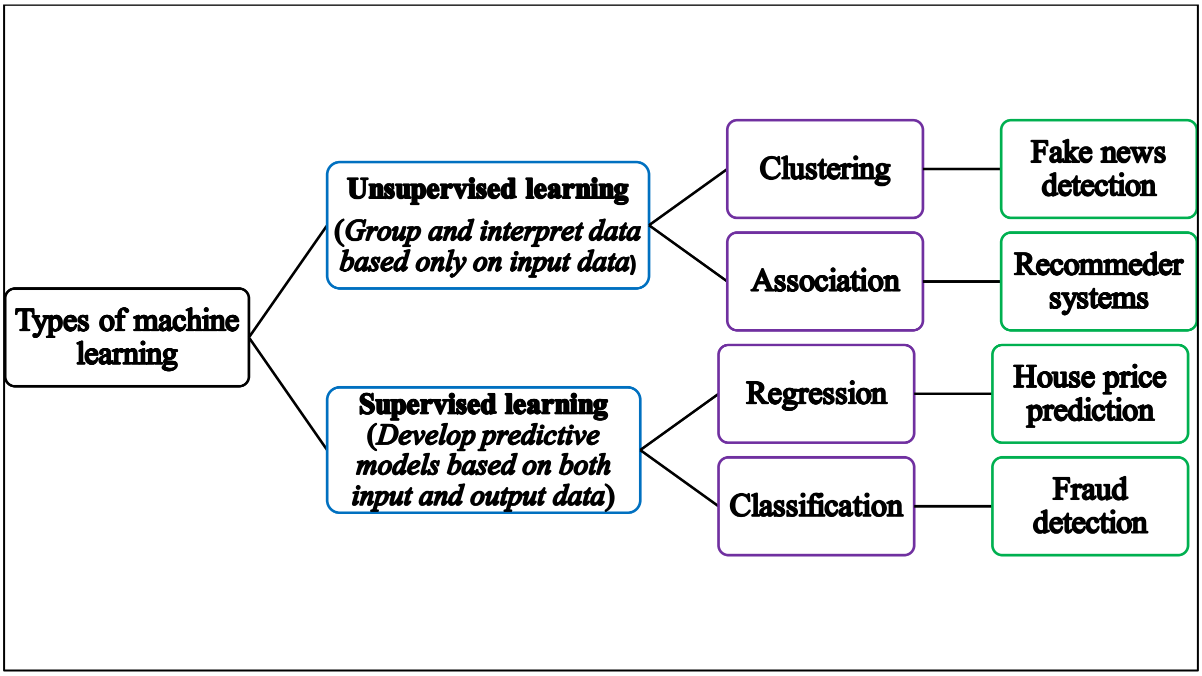

Figure 2.13; Machine learning techniques including both unsupervised and supervised learning. (Adapted from MathWorks, 2020).

2.7.1 Supervised Machine Learning

Normally classification and regression are used for supervised learning to develop models. Some of the common algorithms for performing classification include:

- K-nearest neighbour (kNN)
- Support Vector Machine (SVM)

- Naïve Bayes
- Boosted and bagged decision trees
- Logistic regression
- Neural networks

The following section will give a brief overview of some of the common classification algorithms that will be used for developing the envisioned predictive model discussed in Chapter Six.

2.7.1.1 Supervised Machine Learning Algorithms.

1. **Logistic Regression**

Logistic regression is normally used when solving classification problems. This implies that the target variable is in a categorical format such as yes/no or discrete like a number between 1 and 10. From a given dataset, logistic regression models use an equation to create a curve and then uses this curve to make predictions on new data as shown in Figure 2.14 (Dibble, 2020)

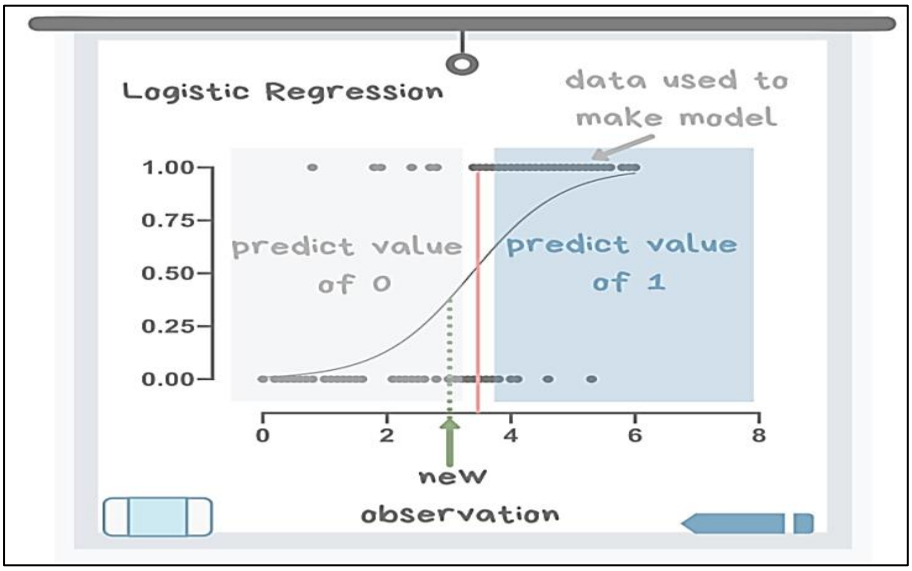

Figure 2.14; Graphical representation of logistic regression (Dibble, 2020).

2. **Linear Regression**

Linear regression uses the best fit concept formulated from modelled data points. This best-fit line is then used to predict values when presented with new observations as shown in Figure 2.15.

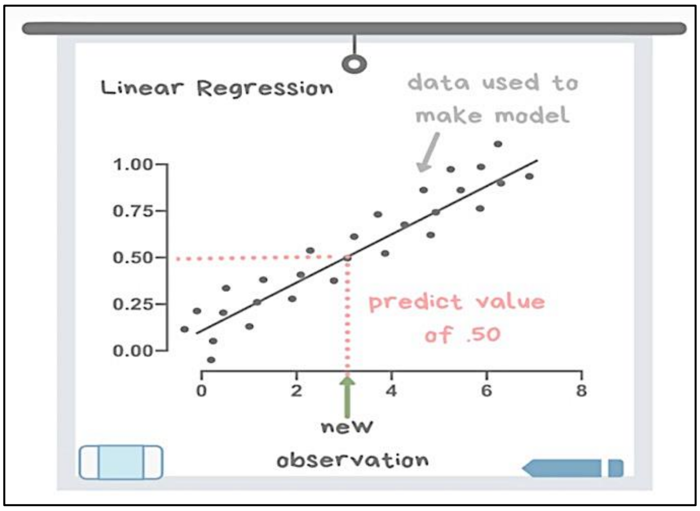

Figure 2.15; Graphical representation of linear regression (Dibble, 2020).

Linear Regression is used for continuous target variables hence essentially can take any value. In fact, any model with a continuous target variable can be categorized as "regression." (Dibble, 2020)

3. K Nearest Neighbors (KNN)

The 'k' in 'kNN' refers to the number of neighbouring points that the model considers when making predictions. The value of 'k' is adjustable until the best results are obtained. Figure 2.16 illustrates the 'k' with reference to new data, the side with more 'k' or influence in the circle determines the classification of the new data. (Dibble, 2020)

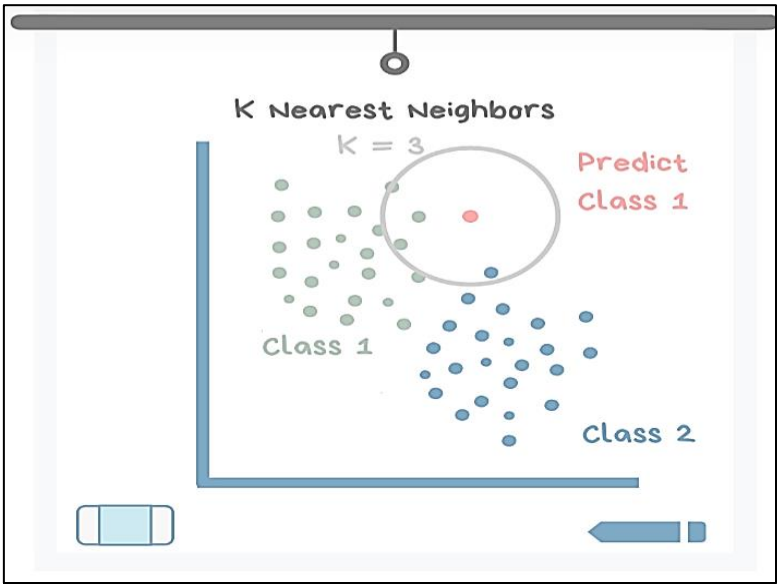

Figure 2.16; Graphical representation of K Nearest Neighbours (Dibble, 2020).

4. Support Vector Machines (SVMs)

SVM algorithms can be used to analyse data for regression and classification problems. These algorithms are binary linear classifiers & non-probabilistic in nature. SVM models map out points so that different categories are clearly separated by a margin/decision boundary that as wide as can be achieved. New data points are then mapped depending on the decision boundary side they fall. Through the Kernel trick, SVMs can also efficiently perform non-linear classification. SVM algorithms construct sets of hyperplanes in a high/infinite-dimensional space. These hyperplanes can be used for classification, regression, and other ML tasks. What makes a good SVM algorithm is the good separation with the largest distance to the functional margins as shown in Figure 2.17. A larger margin will have a lower classifier generalisation error.

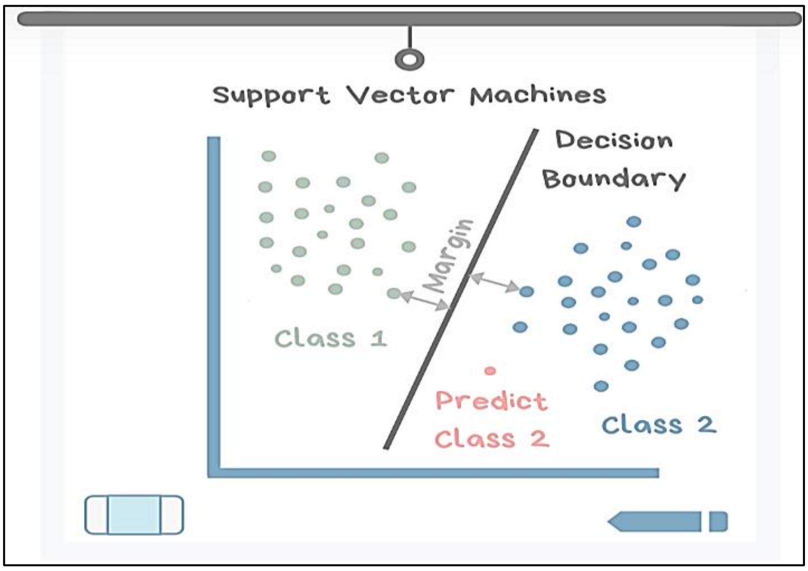

Figure 2,17; Graphical representation of support vector machines (Dibble, 2020).

5. Neural Networks

Neural networks are a series of classification algorithms that identify interactions and interdependencies in a dataset by mimicking the human brain. The “neurons” (the equivalent of a human brain’s nervous system) are mathematical computations that capture and categorise information through a series of layers as illustrated in Figure 2.18. Neural networks are adaptable and sensitive to variations in input hence predict the best possible outcome without having to redesign the output criteria.

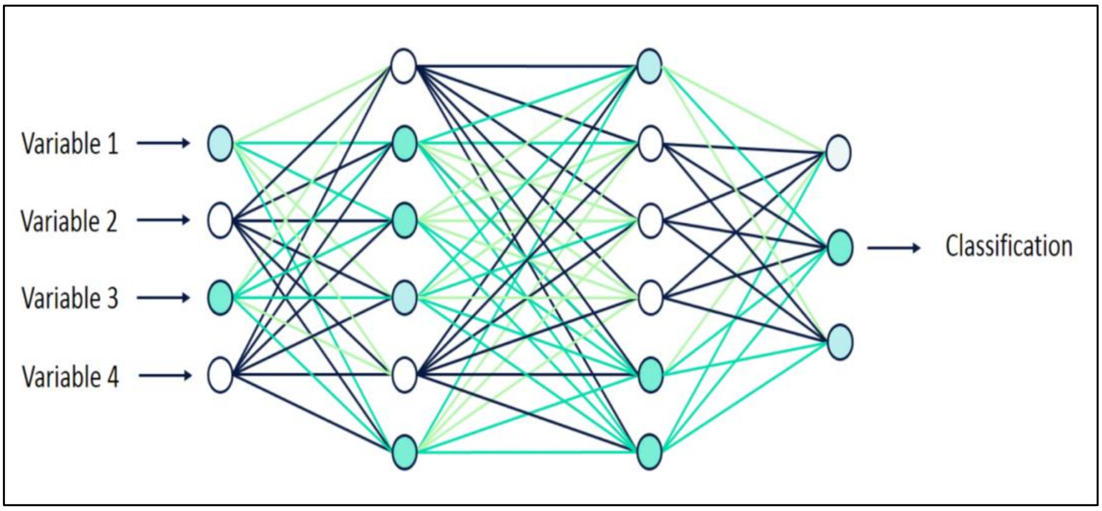

Figure 2.18; Graphical representation of Neural Network layers (Dibble, 2020).

2.7.2 Unsupervised Machine Learning

Unsupervised learning as opposed to supervised, operates such that it learns patterns or other intrinsic features hidden in data. Unsupervised learning is used to analyse input data without tags and draws inferences from the data (MathWorks, 2020) Unsupervised learning is usually used for clustering. For exploratory data analysis, SVM is used to determine if there are hidden features or patterns. Algorithms commonly used for unsupervised learning include

1) K-means
2) Gaussian mixture models,
3) K-medoids
4) Fuzzy c-means clustering

2.8 Chapter Synthesis

Extensive research has been conducted over the years concerning construction field rework due to the negative impacts it poses on overall project performance. The various tools developed in the past have made use of statistical methods to predict the probability of rework occurrence, inter-project learning in an attempt to reduce rework occurrences by learning from past rework events amongst other techniques. In as much as rework reduction programs have been developed to address this challenge, there is a need for further analysis and improvement of these techniques since the problem has continued unabated.

Literature reinforces that elimination/reduction of construction field rework will improve overall project performance and this can be achieved by the early recognition of rework risk triggers to ensure that works are done right the first time. There is great potential for error prediction if sufficient data is gathered and appropriately analysed. Predicting the probability of rework occurrence will act as a project risk management tool that enables proactive control measures to be put in place based on the assumption that if the risk triggers are eliminated or corrected the occurrence of rework will be effectively managed.

Several studies reveal that there is limited research with regards to using technology for quality control during structural concrete construction. Additionally, there is limited research on structural concrete elements in particular. As such, this research seeks to investigate how supervised machine learning can be used as a tool to bridge the gap in knowledge between rework elimination/reduction and predictive data analytics in the construction industry as far as structural concrete elements are concerned. Machine learning is an emerging field and it integrates data-based analytics, data mining, statistics, and pattern recognition to realise interrelationships and trends from various datasets.

Machine learning was selected as a data analytics predictive tool in this research over other data analytics tools such as statistical analytical methods because machine learning algorithms have demonstrated great potential in the ability to resolve engineering problems involving the prediction of failure, defects, and errors. Whenever problem domains are ill-defined and where human knowledge is somewhat limited, machine learning algorithms become very useful. Therefore, building upon knowledge accrued to date, this research aims to provide insights on how best to adopt predictive data analytics in rework management.

2.9 Chapter Summary

The literature review discussed in this chapter provides the reader with a better appreciation of the general definition of rework from a variety of researchers from several countries. This study narrows down rework to construction field related rework with its definition and effects on project performance also identified. It can be concluded from literature that small percentages of rework usually result in excessive losses as far as project investments are concerned. The causes of the identified rework risk triggers are classified to gain a better understanding of the precursors to rework enabling the development of appropriate and effective prevention strategies.

Research has established that the approaches used for quality monitoring in construction sites do not effectively identify any defects early in the construction process. Studies suggest that if quality control is improved during the construction process, the quality of structural concrete also improves resulting in less rework. Several published studies agree that the construction phase of any project is the most critical since nearly 50% of all building failures occur during this period. It is therefore prudent to effectively manage rework occurrence during this phase.

The literature review also provides information regarding the negative impacts of construction related rework on project cost, schedule, and quality performance. It can be noted that due to the high levels of subjectivity in the definitions and interpretations of rework by construction professionals and researchers, standard rework data is usually difficult to obtain resulting in inconsistencies in rework data collection and quantification. However, statistics from literature act as a guide in the estimation of the negative effects of rework on project performance. Rework can account for up to 20% of the project cost, 10 -77% increase in project schedule and 12% of the contract value in the form of quality-related costs.

Available research confirms that even a slight reduction in the cost associated with rework can result in substantial improvements in project performance. Additionally, various tools and strategies to eliminate, minimize and contain rework have been developed over the years as discussed in section 2.6, however, rework remains unabated hence the motivation of this study to identify and develop an efficient data analytics predictive tool that acts as a proactive project risk management strategy in rework elimination.

As discussed in Chapter 2.7, machine learning is an emerging field that can be used in data analytics. Machine learning in this research was adopted to bridge the gap of knowledge between rework management and predictive data analytics. The following Chapter will discuss the tools and techniques used in the collection of rework data from the case study.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Chapter Overview

Discussed in this chapter are the research approach, methods used to collect data as well as the predictive data analytics modelling techniques considered in the accomplishment of the aim and objectives of the research. Primary and secondary data sources which form an integral part of this research are also explained. Section 3.4 and 3.5 review the various tools and accessories used in analysing the data collected from the primary and secondary sources as well as the data mining framework adopted in the development of the predictive model respectively. The development and applicability of the predictive model are extensively discussed in Chapter Six of this study. Illustrated in Figure 3.1 is how this chapter is broken down into seven major sections.

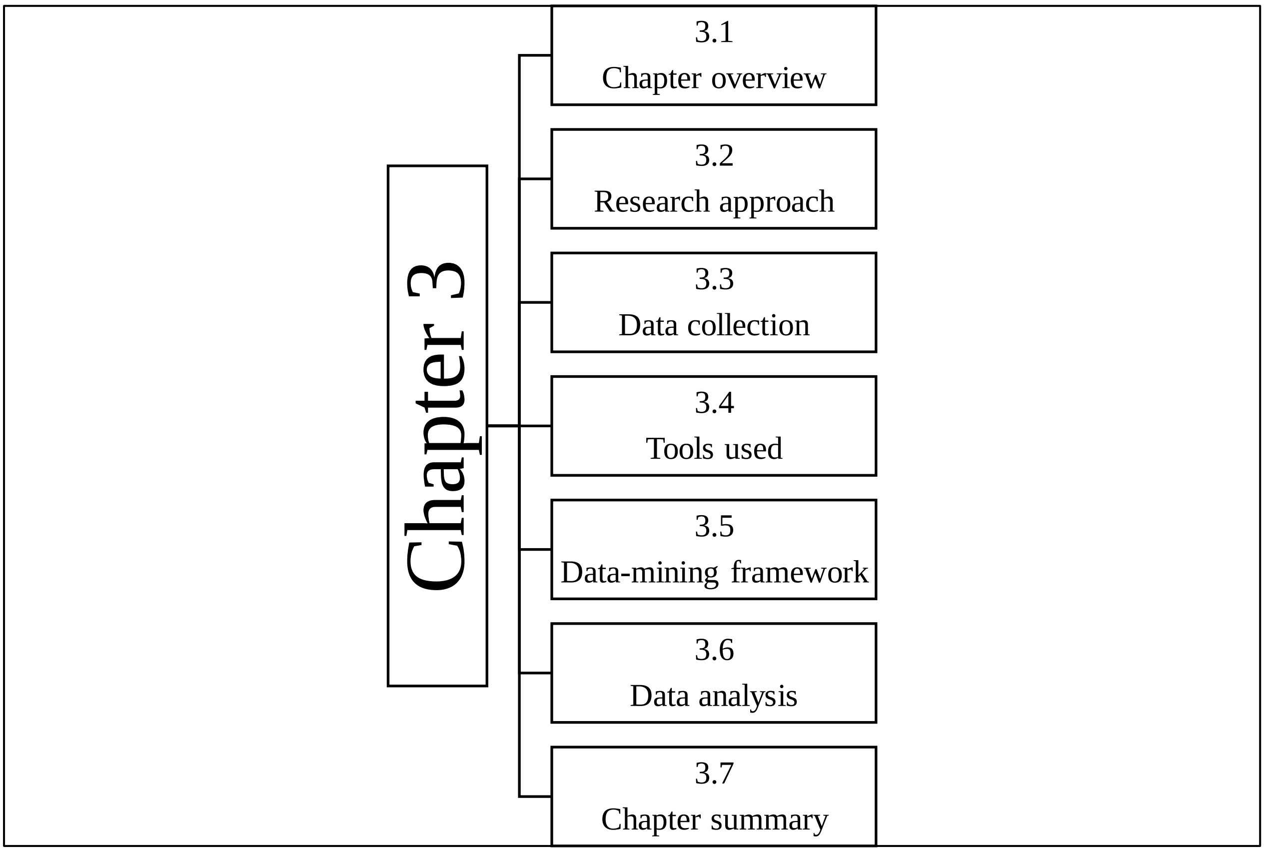

Figure 3.1;Research methodology layout.

3.2 Research approach

The research design is both quantitative and qualitative in nature comprising of an exploratory study, technical site document review, interviews, and a literature review. The research approach focused on the expert interviews (not statistical data) to better understand the phenomena. For this reason, a case study was adopted where construction professionals were interviewed to obtain a better understanding of the process. The case study was also used to collect data and to identify rework risk triggers. According to Leedy and Ormrod (2010), the following fundamental questions surrounding research data need to be answered to collect and formulate the data needed for research:

1) What data was required?

The data required pertains to the various construction risk triggers that lead to reinforced concrete structural element rework. Additionally, preventative and corrective measures need to be identified to proactively minimize the occurrence of construction field rework during the construction phase.

2) Where was the required data located?

The primary data obtained is sourced from a case study (the Hwange Expansion Project construction site) discussed in Chapter 4 of this study). Secondary data will be synthesized from literature of relevant studies.

3) How was the required data be obtained?

The data was obtained through the documentation of regular site inspection observations by the researcher on an online Microsoft Excel Spreadsheet. The analysis and consolidation of the project technical site documentation were also logged on the same spreadsheet. Semi-structured interviews with various experienced construction site professionals were carried out and served as a basis for the formulation of a survey that was conducted on experts in structural concrete construction. Furthermore, relevant data were extracted from an extensive literature study.

4) How was the data collected be interpreted?

Descriptive statistics in the form of statistical commentary, tables, and graphs was used to interpret the data.

3.2.1 Quantitative research

Quantitative research methods are numerical methods used to inquire an identified problem using some testing theory and the analysis and is done by statistical means. They are mainly concerned with the measurement of quantity or the amount and are mostly applied to cases that should be expressed in quantity terms. The main methods which are used to obtain data in quantitative research include:

- Experiments
- Surveys

In this study, a questionnaire survey was used as a quantitative research method for data collection to support the development of the modelling dataset used in Orange (*an open-source software package containing a variety of visualisation tools and classifiers for data mining and analysis*) for modelling.

3.2.2 Qualitative Research

Qualitative research methods primarily focus on the understanding of human and social problems. They involve understanding the reasons for certain behaviours within humans for example their knowledge, actions, beliefs, attitudes, and fears. Reasons for the failure of certain equipment can also be determined using this method of study (Mathi, 2004). The main methods which are used to collect data in qualitative research include:

- Focus groups
- Direct observation
- Interviews
- Questionnaires (Hancock et al, 2009)

Qualitative research methods were employed in this study to determine the variations in the data collected, the criticality of different triggers with respect to the probability of causing reworks, and the general level of experience or level of expertise necessary to effectively undertake each stage.

According to Streefkerk (2020) as summarized in Table 3.1 one of the major differences between qualitative and quantitative research methods is that quantitative methods present data in the form of numbers, graphs, and figures while data in qualitative research is usually presented in non-numerical form as it usually summarizes opinions and concepts or experiences.

Table 3.1; Main differences between qualitative and quantitative research methods (Streefkerk,2020).

	Quantitative Research	**Qualitative Research**
1	Tests theories and hypotheses	Explores ideas and formulates theories and hypotheses
2	Uses mathematics and statistics for data analysis	Summarises, categorises and interprets data
3	Presents data in numbers, graphs and figures	Data is usually presented in words and non-numerical form.
4	Key terms: testing, measurement, objectivity, replicability.	Key terms: understanding, context, complexity, subjectivity.

3.3 Data collection

A hybrid approach was adopted for data collection. Both qualitative and quantitative data collection techniques were used to accomplish the aim and objectives of the research as mentioned in the previous section. This section discusses the collection of primary and secondary data over a twelve-month research period. It should be noted that data collection was disrupted due to the Coronavirus pandemic that instigated site closure and suspension of works for a period of about two months during the second quarter of 2020.

3.3.1 Primary data

Primary data sources involve direct data collection, observations, or measurement of phenomena in the real world. Primary sources of data are also undisturbed by any intermediary interpreter according to Walliman (2005). Additionally, Leedy and Ormrod (2010) are of the opinion that primary data is often the most valid and informative. In this research, primary data was collected in three phases which consisted of routine inspections and observations of the construction project site, audits of site documentary sources such as quality assurance check sheets and non-conformance reports as well as semi-structured interviews with various experienced construction site professionals. The development of the initial research ideas, as well as a more focused and specific research question, can be accomplished by conducting an exploratory study according to Neuman (2000) hence the motivation of the primary data source.

3.4 Tools Used

A variety of tools were used in this study which includes the Orange software programme together with Microsoft Excel. These tools are briefly discussed below.

3.4.1 Orange Software Package

Orange is an open-source software package containing a variety of visualisation tools and classifiers for data mining and analysis. This package was chosen due to its composition that does not require the user to understand classification algorithms from first principles (Botha, 2018). Compared to other data mining resources such as RapidMiner or Scikit-learn, Orange is a free software package where the classification algorithms come in a pre-packaged visual programming interface. Below is an overview of the key features of interest in this study:

(a) Interactive Data Visualization

Data visualization helps uncover hidden patterns in data. It also provides some intuition in data analysis and provides support or communication between various data users. Whilst most visualizations are available in built-in systems, some are available as add-ons and include word clouds, visualizations of networks, geographical maps, and others. (Orange, 2020). Apart from standard visualisation tools found in data mining suites, orange includes extras such as silhouette plot, mosaic and sieve diagram as well as Pythagorean tree.

i. Intelligent Visualizations

When data has multi features, selecting which feature pair to visualize in a scatter plot can be problematic. Intelligent visualisation can be used to solve this feature pair mystery. When provided with class information score plots determine projections with the best class separation. (Orange, 2020). Figure 3.2 illustrates how Orange performs intelligent visualizations.

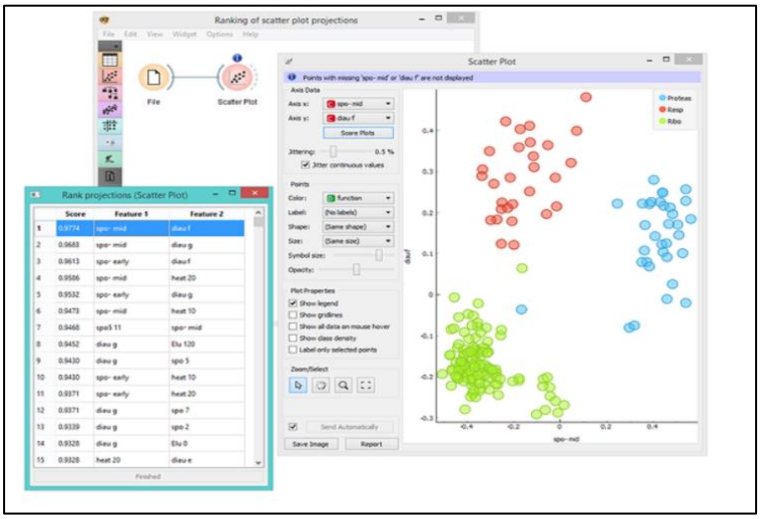

Figure 3.2; Intelligent Visualisation illustration in Orange. (Orange, 2020).

(b) Visual Programming

The following are some of the core components of the visual programming capabilities of Orange:

i. Component-Based Data Mining

In Orange, data analysis is done visually by adding various widgets together. The widgets allow for data retrieval and also for data pre-processing. The software has built-in widgets and some specialised ones can be downloaded as add-ons as shown in Figure 3.3.

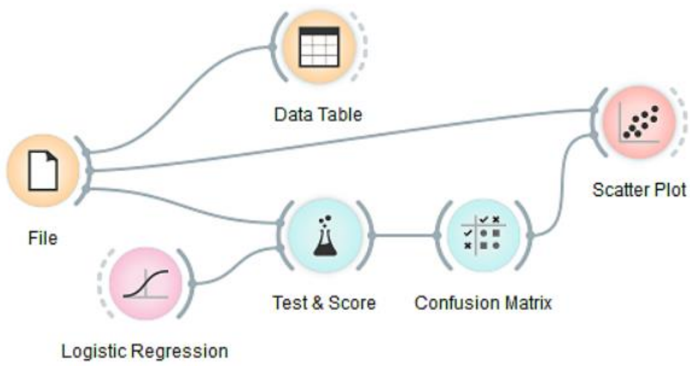

Figure 3.3; Illustration of visual programming widgets added together (Orange, 2020).

3.4.2 Microsoft Excel

Microsoft Excel is a spreadsheet package that contains tools for calculations, graph creation, and a macro programming language called Visual Basic for Applications (VBA) that was used in this research for data collection, analysis, processing, and presentation.

3.5 Data Mining Framework

This study adopted the Cross Industry Standard Process for Data Mining (CRISP-DM) framework for the development of the predictive model. The CRISP-DM framework is a process that covers the whole data science project life cycle. This framework has been frequently used for the data analytics and mining of data science projects (Carneiro da Rocha & Timóteo de Sousa, 2010) and (Wowczko, 2015). This model also acts as a guide to aid in the planning, organizing, and implementation of a data science project. It was applied in this study during the development of the predictive analytics model. Figure 3.4 shows the natural description of the workflow of CRISP-DM in solving data science projects and the stages used to accomplish the aim and objectives of this research.

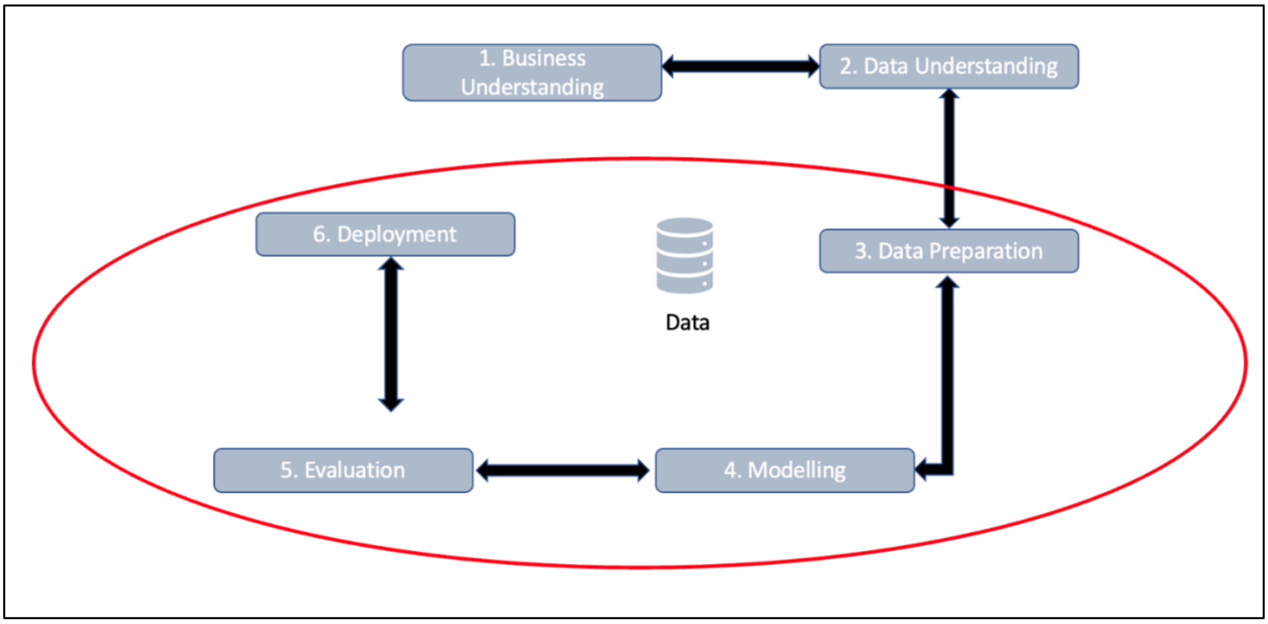

Figure 3.4; CRISP-DM Lifecycle with the last four processses used in this research circled in red (Adapted from Data Science Project Management, 2020).

Steps in implementing CRISP-DM in data science according to Wowczko (2015) include:

1. **Business Understanding:** Aims to appreciate the aim and objectives of the data science project from a business viewpoint. This knowledge is then converted into a data mining problem and a preliminary project plan designed to accomplish the project objectives is formulated thereafter.

2. **Data Understanding:** Focuses on understanding the context of data requirements. It is at this stage that clearly defined data needs and pre-processing requirements are defined to avoid scope creep and possible data management errors.
3. **Data preparation:** Here emphasis is on data representation for modelling. Tools and techniques used for processing the data should be properly understood and the data should be processed to suit the requirements of the software in use.
4. **Modelling:** The techniques suitable for the project should be identified earlier to ensure proper familiarisation of their individual data requirements.
5. **Evaluation:** After identifying several options, model validation and evaluation should be done to determine the best fit model for the organisational objectives.
6. **Deployment:** Depending on the nature of the model, it should be clearly indicated how various stakeholders want to access the results.

3.6 Data Analysis

The data collected from the site inspections, technical documentation, interviews, as well as literature review were analysed as discussed in the following sections.

3.6.1 Site data analysis

The data observed during site inspections as well as that compiled from site documentation was logged on an online Microsoft Excel spreadsheet and analysed by means of content analysis. Trends were also observed and were used to formulate the questionnaire survey as discussed in detail in Chapter 5.

3.6.2 Interview data analysis

The data collected from the various interviews was contextually analysed. The interview notes were transcribed by the researcher and later compared to identify any similarities and contrasts as summarized in Chapter 5. The interview questions are also outlined in Chapter 5.3 of this study.

3.6.3 Questionnaire data analysis

Data from the questionnaire survey was analysed using descriptive statistical methods. Because a Likert 10-point scale was used to rate the influence of the various identified risk triggers, descriptive statistics aided in explaining some of the trends observed as discussed in Chapter

5. The data from these ratings was also used to support the compilation of the dataset that was used for predictive modelling in the study.

3.7 Chapter Summary.

The research methodology chapter gave the reader an insight into how the research was conducted, in terms of the data collection methods, tools, and accessories used together with the descriptive analysis techniques. In essence, this research adopted both qualitative and quantitative research methods with the findings captured in a Microsoft Excel spreadsheet. The source of the primary data (Hwange Expansion Project) is described in the following chapter. The captured data were then pre-processed into a format comparable in the Orange software environment as discussed in detail in Chapter 5 of this study following the CRISP-DM framework. The Orange data mining software package also used in the modelling and evaluation of the machine learning models as detailed in Chapter 6.

CHAPTER FOUR

CASE STUDY OVERVIEW

4.1 Chapter Overview

Hwange Expansion Project, the case study used in this research, is a 660 Mega Watt (MW) coal-fired power plant construction project and this chapter outlines the project background and how the primary data was collected. Also discussed in this chapter are the technical project specifications and requirements in relation to the expected quality of reinforced concrete works. This chapter aims to provide the technical aspects and quality management systems adopted by the Hwange Expansion Project with regards to structural concrete construction. Figure 4.1. illustrates the individual sections contained in this chapter.

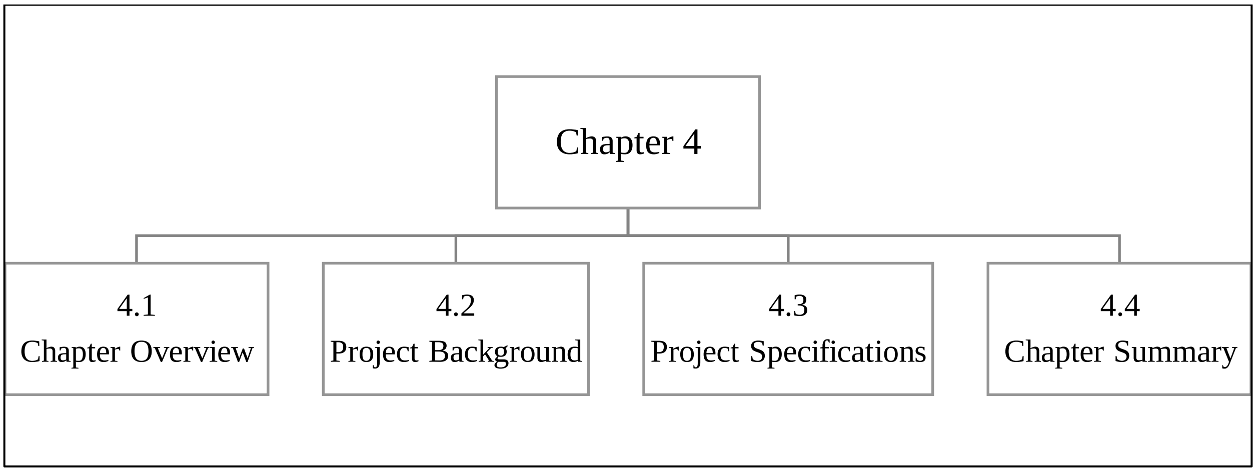

Figure 4.1; Chapter 4 overview.

4.2 Project Background

Hwange coal-fired Thermal Power Station (HTPS) is situated in Hwange town, located in Western Zimbabwe in the Matabeleland North Province, approximately 100 km south-east of Victoria Falls. The HTPS was constructed in two stages with the first stage entering service in 1983 and the second in 1987. Stage 1 comprises 4 × 120 MW turbo generating units while stage 2 has 2 × 220 MW units. The main source of coal to the plant is from the adjacent Hwange Colliery Company Limited (HCCL). Process and cooling water for power plant operations come from the Zambezi River via a 44 km single pipeline. Hwange town, situated around the power station and coal mine, supplies labour and services to these installations. The HTPS

Expansion Project referred to in this research as Hwange Expansion Project (HEP) entails the construction of 2 × 330 MW generation units adjacent to the existing infrastructure of stage 2 units within the HTPS premises as illustrated in Figure 4.2.

Figure 4.2; Illustration of Hwange Expansion Project site relative to the existing Hwange Thermal Power Station.

4.3 Project Specifications

The Hwange Expansion Project (HEP) is a mega power construction project performed under an Engineering, Procurement and Construction (EPC) agreement with a design nameplate capacity of 670MW and is located next to the existing 920 MW installed capacity thermal coal-fired power plant. Hwange Expansion Project is national status granted, with an estimated project cost of United States Dollars (USD) 1.5 billion and expected to be in full commercial operation by the beginning of 2022. Construction of HEP commenced on the 1st of August 2018.

HEP adopts a tailored PRINCE2 project methodology to deliver the engineering, procurement, and construction of a coal-fired plant with two 330MW turbo-generating units, four substations (Hwange B, Lupane, Insukamini and Sherwood) and transmission lines spanning over 360km.

The project infrastructure includes the main power plant structures, local control buildings, *1 x* chimney stack, coal conveyor, handling and storage structures, ash handling structures, *1 x* cooling tower, electrical structures, plant auxiliary structures, administration buildings, and access roads.

4.3.1 Site geology

Based on the geological survey report, the subsoil is mainly made up of jurassic-age volcanic tuff, basalts residual soil, and the triassic-age shale and marlstone. According to the strata structure, distribution, and engineering characteristics, the bearing capacity of the subsoil meets the requirements for power plant construction therefore natural shallow foundations were adopted for some structures. For most structures, the foundation type was concrete isolated foundation or concrete block foundation. The foundation depths ranged from 0.50 to 5.00m. The groundwater levels also ranged from 0.50 to 11.50m depending on topography. The geotechnical investigation revealed that the groundwater had a weakly corrosive effect on the reinforced concrete structures.

4.3.2 Reinforcement steel specifications

The steel reinforcement for the concrete structures was chosen based on the SANS 920:2011 and BS 4449:2005 specifications. For the yield strength of 250 MPa, the 6mm, 8mm, and 10mm nominal diameter steel reinforcement bars were adopted while for 450Mpa yield strength the 12mm, 16mm, 20mm, 25mm, 32mm, and 40mm nominal diameter steel reinforcement bars were used. Both lap or mechanical splices were used in this project depending on the design component.

An inspection check sheet for steel reinforcement (Attached in Appendix A) specified the inspection of the following design parameters:

1) Tensile strength
2) Surface quality
3) Rebar spacing tolerance
4) Rebar size
5) Concrete cover
6) Overlap length
7) Tie tightness
8) Link spacing

9) Joint positioning

The tensile strength of the reinforcement bars was verified at the site laboratory and depended on the bar size as previously mentioned. The rebar spacing tolerance was 5mm. The rebar size, overlap length, link spacing, and cover to reinforcement were design and structural element specific. All the ties had to be sufficiently tight to ensure that rebars were not displaced during concrete placement. For jointing, mechanical splices had to be placed in a staggered manner to disrupt the possible line of weakness. The minimum binding lap length was 50d where (d) is the diameter of the lapping rebars. The surfaces of the steel rebars were visually inspected and had to be straight, without cracks, grease, grains, or flaky rust.

4.3.3 Formwork specifications

The project used both steel and plywood formwork for concrete casting with vertical formwork systems used for columns and horizontal systems used for slabs and beams. Steel formwork was adopted in instances of structural elements with longer spans or elements with particular shapes such as the herringbone columns of the cooling tower. The formwork was removed from the concrete mould after sufficient curing was achieved. The formwork inspection check sheet (Attached in Appendix B) listed the following parameters to inspect:

1) Final elevation
2) Formwork seam
3) Axis displacement
4) Surface quality
5) Formwork type
6) Adequacy of concrete spacers
7) Formwork bracing

Both the final elevation and axial displacement were measured using a total station and the allowable deviation in both instances was 5mm. The surface quality of the formwork during inspections had to be clean, smooth, level, and sufficiently lubricated. Formwork seams had to be sealed to prevent grout loss that can potentially lead to honeycombing.

4.3.4 Reinforced concrete specifications

Subject to the structural design, the concrete strengths adopted for the research study varied from C15 (15MPa) to C40 (40Mpa). Concrete was batched and mixed at a batching plant 1km away from the construction site and was transported to the site by means of concrete mixers.

Concrete inspection checks (Attached in Appendix D) were done on fresh and hardened concrete. During the placement of concrete, the following parameters were inspected:

1) Mix design
2) Target slump
3) Vibration
4) Target temperature

The concrete mix design was determined at the batching plant before delivery to the site. During placement, the slump was determined to ensure that it was within the design range. Slump ranges varied from structure to structure but were usually 110 – 190 mm for C30 slabs. Depending on the accessibility and packing of reinforcement bars, 50mm diameter soft shaft poker vibrators were used for vibrating the concrete mix but different shaft diameters were also used. The target temperature during placement was also measured before a pour and the expected concrete temperature range was 5 to 35 ^{o}C. The check sheet used for slump inspections is Attached in Appendix C of this report.

After the placement of concrete, curing of the element was done for not less than 7 days. The formwork was removed after 7 days to facilitate cast concrete inspections. The following parameters were inspected after the end of the curing period and formwork is uninstalled:

1) Elevation
2) Curing period
3) Surface quality

The final elevation of the cast concrete element was determined using a total station and the acceptable deviation was ±30 mm. The surface quality of the cast concrete had to be smooth and free from any cracks, defects, and honeycombing.

4.4 Primary data collected

4.4.1 Routine site inspections

The researcher conducted regular site inspections weekly over a period of twelve months as an exploratory research technique and recorded rework observations on an online site inspection Microsoft Excel spreadsheet. The routine inspections provided a basis for the formulation of the questionnaire survey which was used as input for modelling purposes. The routine inspections were conducted pre, during, and post concrete placement. Routine site interviews were also conducted following a rework incident to establish the underlying risk triggers that would have contributed to the rework occurrence. The routine inspections were conducted in phases, and various checks were made at each phase as summarized in Figure 4.3.

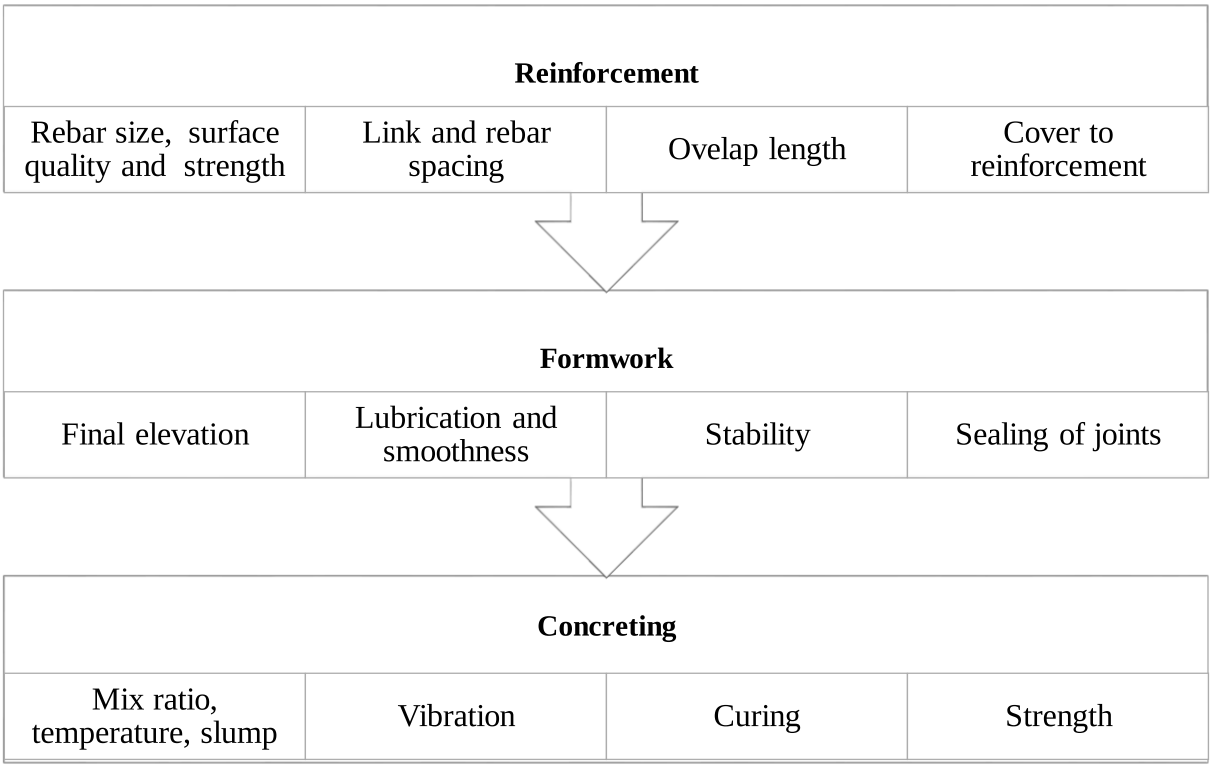

Figure 4.3; Routine inspections made by the researcher on site.

Some of the semi-structured interviews followed a rework incident observation to establish the underlying risk triggers that would have contributed to the rework occurrence.

4.4.2 Technical site document audit

The Hwange Expansion Project adopts a tailored Projects in Controlled Environment (PRINCE2) project management methodology and implements quality assurance arrangements in compliance with the requirements stipulated in International Organisation for Standardization (ISO) 9001:2015. The Project Quality Management Plan required that all tests and inspections conducted be documented. For the purposes of this study, rework data was compiled from:

1) Non-conformance reports
2) Quality inspection sheets
3) Corrective action reports
4) Hazard Identification and risk assessment reports

These site documents confirmed that activities had been executed or results had been attained. Records can, for example, be used to show transparency that stipulated assurance standards and requirements are being observed and met, verification is being performed and that preventative and corrective measures are being implemented. The rework data from these records was used to populate the online site inspection Microsoft Excel spreadsheet.

4.4.3 Interviews

Interviews are defined as distinctive research techniques which involve discussions between two people with the intent of obtaining an individual's perspective on a particular subject matter (Hancock et al, 2009). Furthermore, an interview provides extensive insights into the respondent's opinion which can provide the basis for the identification of variables and relationships.

In this research, semi-structured interviews with various construction site-based professionals who deal with structural concrete construction daily were conducted as an exploratory research technique to establish the construction risk triggers that lead to rework in reinforced concrete construction in the Zimbabwean construction industry. The interviews also served as a basis for the formulation of a questionnaire survey that was conducted on a variety of construction professionals.

To obtain diverse and unbiased responses, the semi-structured interviews were conducted on a variety of professionals which included site agents and engineers, forepersons as well as Safety Health Environment Risk and Quality (SHERQ) practitioners with varying years of experience, professional background, and academic qualifications. The summary of data collected from interviews is extensively discussed in Chapter 5 of this research. Some of the interviews were conducted in person and some on virtual platforms ranging from 10 to 30 minutes with the intent to:

1. Identify the construction risk triggers that result in rework during the installation of steel rebars and formwork as well as during concrete placement.

2. Identify any preventative measures that can be taken to prevent the occurrence of these risk triggers.

4.4.4 Questionnaire Survey

A questionnaire survey enables the collection of data beyond the physical reach of the researcher. Therefore, it can be established that a questionnaire is an impersonal probe. Due to the level of impersonality associated with questionnaire surveys, they need to be administered by realistic standards according to (Leedy and Ormrod, 2010). The grammar and vocabulary used must be unambiguous and unmistakably clear for the benefit of the respondent.

Additionally, questionnaires are particularly designed to investigate a specific research objective, as questions are often inexpertly written, and this results in a low response rate (Leedy and Ormrod, 2010). Furthermore, questionnaire surveys should not entail extensively complicated questions to facilitate truthful responses. The questionnaire in this research targeted a variety of construction experts involved in structural concrete construction in the Zimbabwean context.

Fellows and Liu (2008) suggest that questionnaire surveys be distributed by post or via the internet to respondents. The questionnaires were administered via email due to the Covid-19 pandemic restrictions. The structure of the questionnaire aimed to determine to what extent the identified risk triggers contribute to structural concrete rework. The questions for the survey were compiled according to the findings of the literature study, interviews, and site inspections. The questionnaire comprised of two main sections, namely:

1) Profile of respondents
2) Effects of identified rework triggers on structural concrete rework

The first section of the questionnaire required some background information regarding the profile of respondents which included their professional background, years of industrial experience, as well as the highest level of education. A summary of the findings of the questionnaire survey is discussed later in detail in Chapter 5. The various construction professionals rated the influence the identified triggers have on structural concrete rework using a ten-point Likert-scale in the last section. The questionnaire survey is attached in Appendix G of this study.

4.5 Chapter Summary

The aim of this chapter was to provide the reader an insight into the source of primary data used in this research. The background to the Hwange Expansion Project together with its motivation and description is also presented. Hwange Expansion Project implements quality management systems that meet the requirements set out in ISO 9001:2015. The Project Quality Management Plan stipulates the use of various local and internationally recognised standards for quality control and assurance. According to the employer's requirements, all civil works must comply with the recognised codes of practice and design codes applicable to the works. Some of the technical documents used in the study's investigation included various inspection check sheets derived from the approved standards and codes of practice.

CHAPTER FIVE

DATA ANALYSIS AND PROCESSING

5.1 Chapter Overview

A review of the analysis of the primary data collected by means of an exploratory case study, technical site document review, expert interviews as well as a questionnaire survey is discussed in this Chapter. Descriptive statistical methods were employed to describe and meaningfully summarize insights from the data collected. The interpretation of the analysed results from various data sources is also discussed.

In order for a predictive model to learn, a suitable dataset should be provided. Because datasets relating to structural concrete rework elements are limited and not easily accessible, this chapter discusses the type of data that can be captured and how it can be used to compile a dataset that can be used for predictive modelling.

The preparation and processing of the dataset presented in section 5.5 of this chapter describes the various steps and logic behind the formulation of the fictitious dataset from the data collected in the questionnaire survey. It should be noted that these steps were taken to illustrate how data can be used in predictive modelling. Some of the tools and accessories introduced in Chapter 3 are extensively used in this chapter for data analysis and processing. Figure 5.1 summarises how the various data sources were integrated into the formulation of the final dataset for modelling.

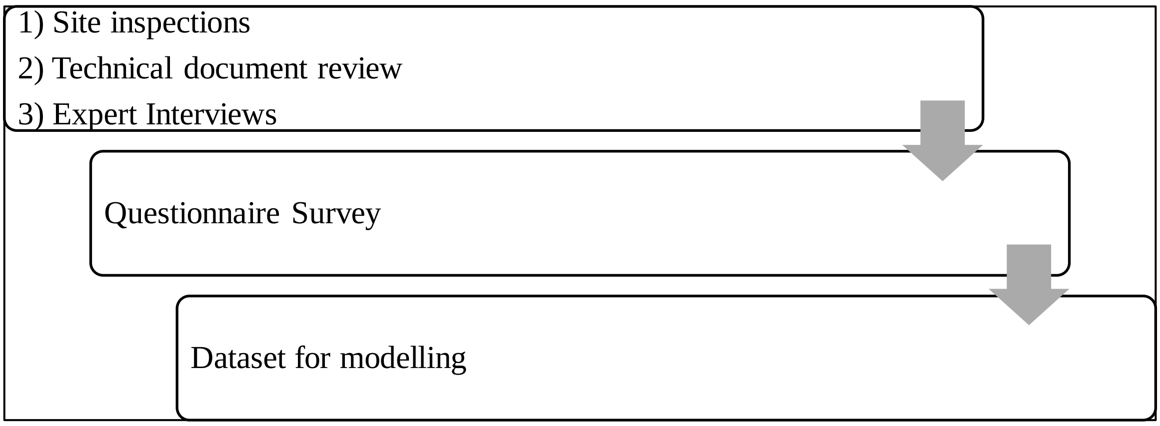

Figure 5.1; Integration of the various data sources used in the development of the dataset used for predictive modelling.

5.1.1 Definition of terms

1. **Attributes** – Used interchangeably with the term "variables" as they refer to the column inputs (rework triggers) in the dataset.
2. **Dataset** – A collection of data presented in a tabular form consisting of columns (attributes/ variables) and rows (instances).
3. **Instances** – Correspond to the number of rows in the dataset, in this case, represents the number of respondents in the questionnaire survey.
4. **Variables** - Used interchangeably with the term "attributes" as they refer to the column inputs of the dataset.

5.2 Site Observations and technical documents review

Chapter 3 describes how the researcher carried out routine site inspections over a period of twelve months at the Hwange Expansion Project construction site. Observations from these investigations formed a fundamental role in the development of the research and also contributed to the formulation of a questionnaire survey that was distributed among various construction professionals. The findings relevant to this study were logged on an online Microsoft Excel Spreadsheet as mentioned previously.

5.2.1 Site observations

The findings were categorised into the three concrete construction phases as discussed in the previous chapters (preplacement, placement, and post placement). Site observations commenced in October 2019 soon after attainment of ethics clearance from the University of Stellenbosch with the reference number ING-2019-11592. Site observations were carried out on a weekly basis until mid-December 2020 (Christmas break).

Hwange recorded flash floods in January and February 2020 which resulted in the suspension of some works at the construction site due to the heavy rains. Following the flash floods, the Covid-19 pandemic also disrupted data collection as the construction site was temporarily closed from March to May 2020 as per mandate from the Government of Zimbabwe. The researcher regained full access to the site end of May 2020 as the Government of Zimbabwe eased lockdown restrictions to essential service providers.

During reinforcement inspections for various structural elements, the following parameters adopted from inspection check sheets and literature were logged:

1. Type of structural element
2. Date, day, and time of inspection
3. Distance from the site office and location of the inspected element.
4. Type of inspection conducted (cover to reinforcement, rebar spacing, etc)
5. Findings from the inspection

Due to the flash floods experienced in January of 2020, most reinforcement inspections carried out during or after the period had a large number of rusted rebars.

As far as formwork inspections were concerned, the researcher noted the following in the log:

1. Type of structural element
2. Date, day, and time of inspection
3. Type of formwork
4. Distance from the site office and location of inspected element.
5. Type of inspection conducted (elevation checks, formwork bracing and support, etc)
6. Findings from the inspection

This project made use of plywood formwork for the majority of the works since most of the structural elements were of regular shapes. One interesting element was the herringbone column of the cooling tower that required fabricated steel forms due to its irregular shape.

For concreting works, both during and post placement, the following items were logged:

1. Type of structural element
2. Date, day, and time of inspection
3. Distance from the site office and location of inspected element.
4. Volume of concrete cast
5. Temperature and slump recorded
6. Curing method

Hwange is considered an arid region with an average annual temperature of about 35^{o}C therefore concrete placement works were usually carried out early in the mornings or later during the day at cooler temperatures. The immersion method was the most used curing method although some structures such as the chimney stack and cooling tower made use of curing agents.

5.2.2 Site documents

Various technical site documents were reviewed to better understand the underlying factors that contribute to structural concrete rework. The following documents were studied, and the findings contributed to the formulation of the questionnaire survey:

1) Non-conformance reports
2) Quality inspection sheets
3) Corrective action reports
4) Hazard Identification and risk assessment reports

These technical documents provided insight into the root causes of various rework incidences. The Non-conformance report (Attached in Appendix E) for instance, required details of the findings, an investigation of root causes as well as proposed corrective measures that can be taken to avoid reoccurrence of a non-conformance. The Hazard Identification and risk assessment reports (Attached in Appendix F) were developed at every construction stage before the commencement of works. This exercise was a preventative measure in which professionals would discuss the potential hazards and risks associated with a construction activity and suggest measures that can be adopted to mitigate against them. The other quality assurance check sheets are attached in the Appendices section as previously discussed in Chapter 4.

5.3 Interviews

Interviews with various construction professionals were carried out as a basis for the design of the questionnaire survey as mentioned in Chapter 3. This exercise together with site inspections and technical document reviews was conducted throughout the research. These activities were considered an exploratory study since they aimed at attaining a deeper understanding of the various rework risk triggers that cause structural concrete rework during construction.

The interview participants were identified through Hwange Expansion Project Human Resources records as they were employees of the construction project. Seven experts of varying professional backgrounds and academic qualifications were interviewed. Table 5.1 provides a summary of the interviewee profiles. Some interviews were conducted virtually due to the outbreak of the Covid-19 pandemic therefore, prior to the virtual interview, the respondents were contacted via telephone to confirm their availability. The same code of discussion was also used for face-to-face interviews.

Table 5.1; Summary of interviewee profiles.

Participant	Interview type	Current title	Highest level of education	Experience in concrete construction (years)
1	Face to face	Section Engineer	Master's Degree	10
2	Online	Senior Civil Engineer	Master's Degree	17
3	Online	Systems Engineer	Honours Degree	9
4	Online	Project Engineer	Honours Degree	5
5	Face to face	Civil Engineer	Master's Degree	11
6	Online	Civil technician	Diploma	12
7	Online	Quality Engineer	Honours Degree	7

The interview guide comprised of three sections. The first interview section collected the profiles of the participants as summarized in Table 5.1. It can be concluded that a diverse group of experts of varying professional backgrounds, levels of education, and experience in the industry were interviewed.

The second interview section was semi-structured with the intent of gathering as much information as possible regarding other potential risk triggers the researcher might have overlooked during site visits and technical document analysis. The researcher discussed the identified risk triggers per category and invited the participants to add other triggers that could have been missed. The following section summarizes the responses to the question which was:

"Are there any other risk triggers in the three phases involved in the construction of reinforced concrete structural elements you would like to add?"

1) Reinforcement

- Inadequate starter lengths

- Incorrect bending of rebars
- Mechanical damage to reinforcement e.g., when people walk over slab mesh

2) Formwork

- Poor workmanship from unskilled personnel

3) Concrete works (during placement)

- Concrete mix stored for considerable periods of time in a mixer before placement resulting in segregation due to overmixing
- Poorly graded or wrongly sized aggregates
- Poor/wrong mix design e.g., insufficient cement
- Placement of concrete exceeding minimum height requirements (e.g., pouring concrete over 2m heights).
- Mechanical damage to finished concrete surface e.g., personnel walking on a freshly poured slab
- Cold joint formation due to unavailability of onsite batch plant or interruptions in concrete delivery to site
- Delayed delivery of other materials
- Crowded reinforcement areas leaving no space for uniform placement of concrete (e.g., joint of a column and a beam and/any other structural member)

4) Concrete works (post placement)

- Removal of formwork before sufficient strength is gained

A considerable number of rework risk triggers were identified especially during concreting works.

The last interview section required the participants to discuss any rework incident they have come across and the corrective/preventative measures that were taken thereafter. Table 5.2 summarizes the various rework incidence encountered by the interviewees.

Table 5.2; Rework incidences encountered by interviewees.

Rework Incident	Structural element	Construction stage when incident was observed	Corrective/preventative measure taken
Honey combing	Stub column	Post placement	Grouting using epoxy grouting compound
	Beam	Post placement	Plaster of honeycombs using cement grout
	Column	Post placement	Surface grinding and plastering
Formwork collapse	Column	During placement	Bracing redone. Use of metal formwork
Shrinkage cracks	Slab	Post placement	Pouring concrete during the cooler parts of the day e.g. early in the morning or at night
	Beam	Post placement	Use of grout to fill the cracks
	Beam	Post placement	Ensure adequate curing takes place.
Deformed element	Beam	Post placement	Demolition
Formwork displacement	Column	During placement	Stub redo
Axis displacement of the concrete structure which is noticed after the formwork has been stripped off	Column	Post placement	Demolish disfigured element-scarify connection point-apply bonding agent-corrective pour
Damaged concrete surface after formwork removal	Wall	Post placement	Apply thick plastering to correct the bulged wall.
Formwork collapse	Column	During placement	Redoing the shuttering and properly bracing it
	Beam	During placement	Use of proper vibrators

Rework Incident	Structural element	Construction stage when incident was observed	Corrective/preventative measure taken
Formwork burst due to over poking and poorly braced formwork	Column	During placement	Vibrate concrete properly and always have a standby poker on site during pouring
Misalignment of a stub column due to formwork displacement	Column	Post placement	Proper chronology in setting out checks from Total Station. By artisans, foreman, site engineer and then client resident engineer before concrete placement.
Low 28-day strength	Beam	Post placement	Demolition

Most of the structural reworks were discovered post placement, emphasising the importance of thorough quality assurance exercises prior to the placement of concrete. As a preventative measure, one interviewee suggested conducting inductions and pre task risk assessments with the workforce prior to the commencement of any construction work.

5.4 Questionnaire Survey

An online questionnaire survey designed via Google Forms was distributed to several professionals at the Hwange Expansion Project who were identified and invited to participate due to their experience in structural concrete construction. The survey aimed to establish the extent to which the identified risk triggers influence structural concrete rework. The survey dataset was later processed for modelling. It should be noted that the survey results in this study were only used to formulate and assess a process to develop models that can be implemented in real life projects. In a real project setup, actual rework incidences are to be recorded to capture sufficient data that can be used for modelling purposes.

Participants were sent the survey link via personal email after a confirmation phone call. This section discusses the analysis of the data gathered in the survey using the questionnaire. The survey complied with the requirements of the Stellenbosch University Ethics Approval Committee with the reference number ING-2019-11592. The survey questions are attached in Appendix G. The participants were provided with a brief background to the study to appreciate the nature of the questions together with the aim and objectives of the research.

A total of 33 participants responded to the online questionnaire. Figure 5.2 summarizes the professional backgrounds of the various participants and it can be observed that over 60% of the respondents are site engineers.

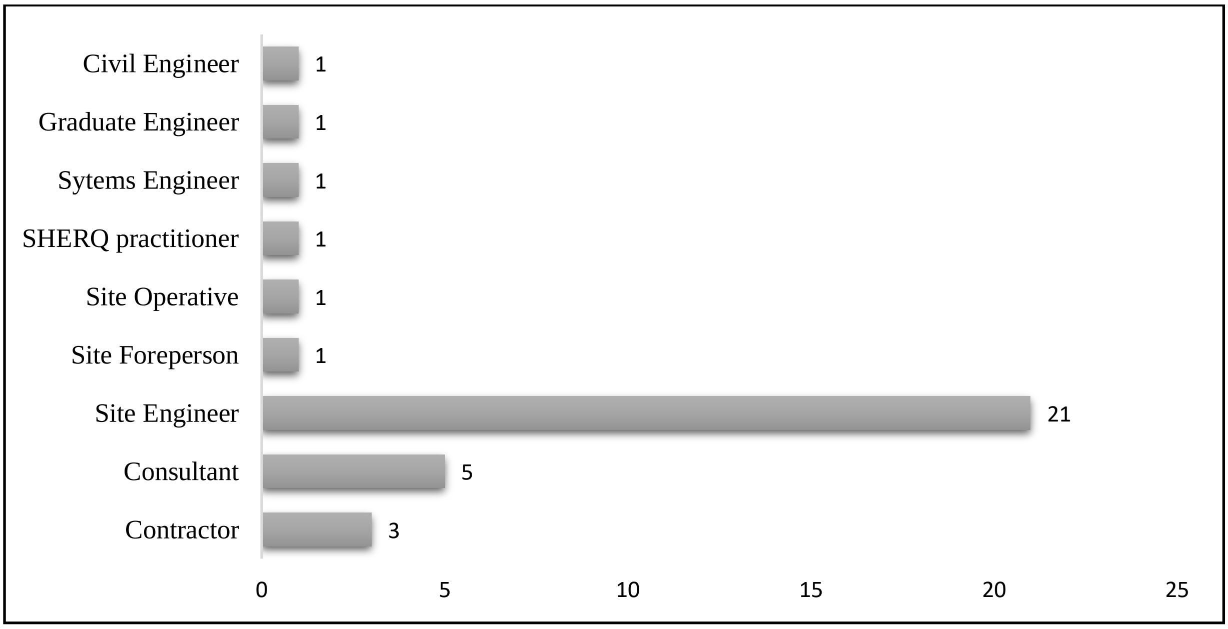

Figure 5.2; Professional backgrounds of respondents.

Of the 33 respondents, over 60% have an Honours degree as the highest level of qualification as shown in Figure 5.3.

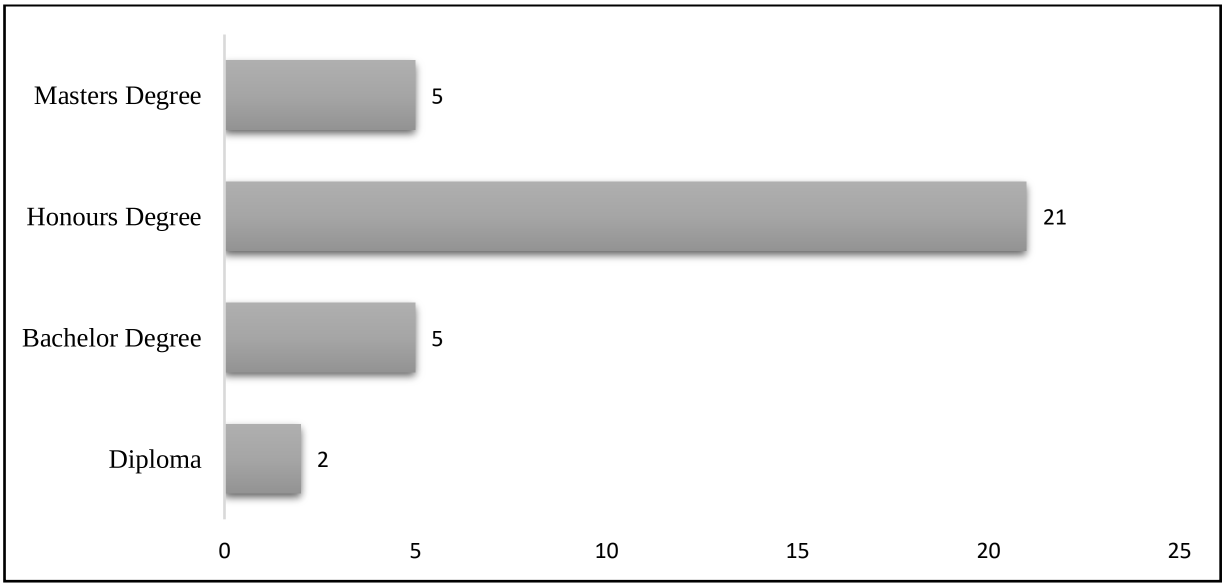

Figure 5.3; Respondents highest level of academic qualifications.

Approximately 80% of the respondents have up to 10 years of construction experience as shown in Figure 5.4. 15% had the experience of up to 20 years and around 6% have over 20 years of experience in the construction industry.

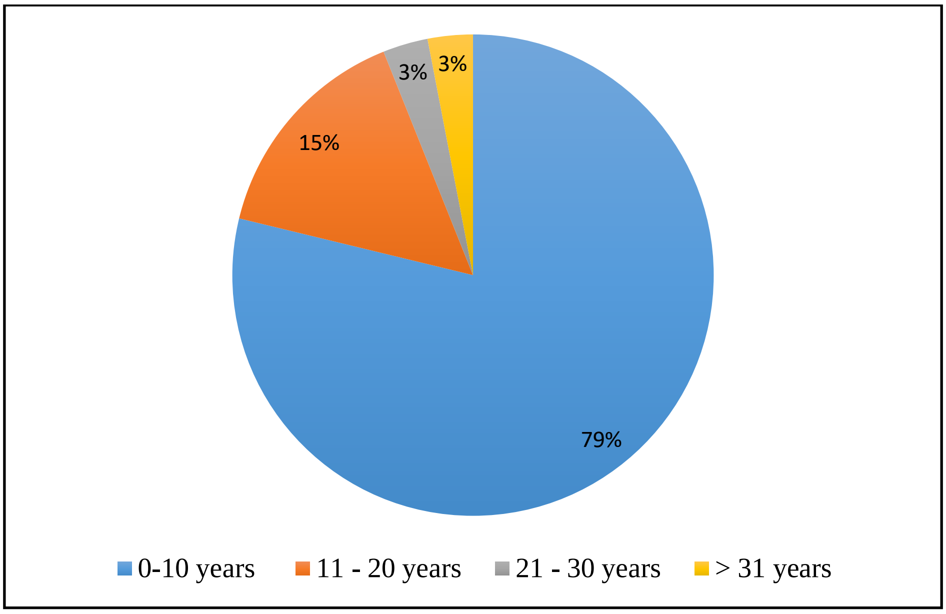

Figure 5.4; Respondents years of construction experience.

As mentioned above, the survey was conducted to establish the extent to which the identified risk triggers influence structural concrete rework at each construction stage. Reinforcement, which was established as the first stage earlier in the study comprised of a total of 9 risk triggers. Table 5.3 summarizes how the various risk triggers identified during the placement of reinforcement influence structural concrete rework. The respondents ranked missing reinforcement bars as the greatest influencer as can be seen by its mode, average and total rating. The average is the mean of the rating of all the respondents whilst the mode represents the score that the most participants rated. Table 5.3 also shows that the greater the average rating, the more the influence of that particular trigger on structural concrete rework.

Table 5.3; Results from reinforcement rework risk triggers.

Rework Risk Trigger	Rating					Ranking
	Min	Max	Mode	Av	Total	
Missing reinforcement bars	1	10	10	8	262	1
Incorrect rebar positioning and placement	3	10	10	7	243	2
Incorrect cover to reinforcement	3	10	8	7	242	3
Incorrect rebar sizes	2	10	8	7	241	4
Inadequate lap distance	3	10	10	7	239	5
Insufficient rebar tensile strength	1	10	10	7	236	6
Poor rebar surface quality	1	10	5	6	205	7
Loose ties	1	10	8	6	200	8
Un-staggered mechanical splices and connections	0	10	5	6	190	9

Similar to the findings by Manuel, et al., (2018), insufficient formwork support and bracing were found to be the greatest rework influencer as far as formwork is concerned as shown in Table 5.4. It can be noted from the table that insufficient support and bracing had the highest mode, average and total rating. Wrongly dimensioned formwork together with poorly sealed formwork joints also had high ratings indicating that these triggers also greatly influence rework occurrence.

Table 5.4; Results from formwork risk triggers.

Rework Risk Trigger	Rating					Ranking
	Min	Max	Mode	Av	Total	
Insufficient formwork support and bracing	1	10	10	8	265	1
Wrongly dimensioned formwork	1	10	10	8	255	2
Inadequately sealed formwork joints	1	10	10	7	246	3
Deformed or damaged formwork boards	2	10	9	7	234	4
Incorrect formwork elevation	1	10	10	7	231	5
Poor surface quality of formwork board	2	10	8	6	199	6

During concrete placement, it was observed that inadequate vibration closely followed by formwork displacement and placement of concrete under adverse weather conditions greatly influence concrete rework as summarized in Table 5.5. This corresponds to the discussions raised in Chapter 2 of the study.

Table 5.5;Results from concreting risk triggers.

Rework Risk Trigger	**Rating**					**Ranking**
	Min	Max	Mode	Av	Total	
Inadequate vibration	4	10	7	8	257	1
Formwork displacement during placement	1	10	10	8	254	2
Placement of concrete under adverse weather conditions	3	10	6	8	252	3
Poor concrete workability	0	10	8	7	242	4
Grout loss	2	10	8	7	222	5
Erratic delivery of concrete mix to site	1	10	6	7	221	6
Over workability of concrete	1	10	8	6	214	7
Using inappropriate poker head for vibrating	0	10	8	6	207	8
Over vibration	0	10	8	6	193	9

Summarized in Table 5.6 are the responses from the two risk triggers identified during post placement. Poor and inadequate curing was regarded as the greatest trigger for rework relative to mechanical damage during stripping. It should be noted that, because there were only two variables in the post placement phase, these triggers were not considered in the development of the predictive model which is discussed in detail in Chapter 6. The following section gives a brief overview of the steps taken in transforming the data collected from the questionnaire survey to a dataset suitable for modelling purposes.

Table 5.6;Results from post concreting risk triggers.

Rework Risk Trigger	Rating					Ranking
	Min	Max	Mode	Av	Total	
Poor and inadequate curing	1	10	7	7	227	1
Damage during stripping	1	10	2	5	169	2

5.5 Data preparation

Data from the questionnaire survey was captured in Microsoft Excel in a format that resembles an expert knowledge system. An expert system is a computer-based program that mimics human expert knowledge of a particular field of specialty. Normally, expert systems are used for repetitive human expert knowledge-based tasks where an actual human expert is not available or cannot be made available. The data collected during investigations was processed to reflect expert knowledge of the various professionals who are involved in structural concrete works. The knowledge was mainly based on answering objectives 1 and 2 which are:

1. To identify construction risk triggers that lead to structural concrete rework during the project construction phase.
2. To categorize the identified risk triggers into phases during reinforced concrete structure construction.

The expert system approach which is illustrated in Figure 5.5 was adopted in an attempt to answer objectives 3 and 4 which are:

3. To develop a suitable dataset that can be used to train and test the performance of classification algorithms.

4. To apply predictive data analytics in developing a model that can be used to reduce the occurrence of construction field rework during construction.

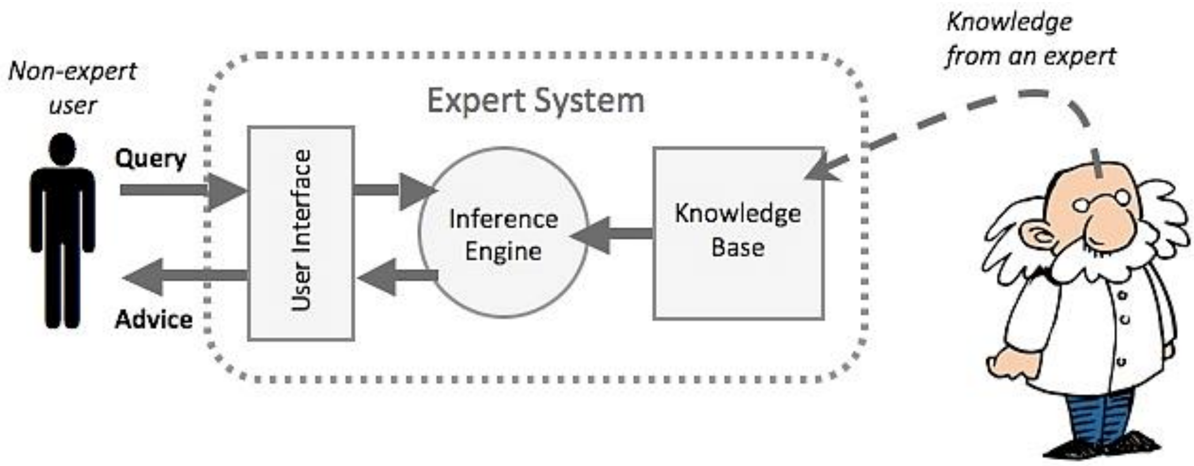

Figure5.5;Expert system approach used in the research (https://www.igcseict.info/theory/72/expert/index.htm).

The initial dataset as illustrated in Table 5.7 comprised of 26 columns (variables) which represent the identified structural concrete rework risk triggers. The dataset also consisted of 33 rows which correspond to the 33 survey participants with each value in a row representing participant rating of how much each rework risk trigger influences structural concrete rework in the categories of the placement reinforcement bars, formwork, and fresh concrete.

Table 5.7; An extract of the initial dataset formulated from the questionnaire survey responses.

Respondant	Insufficient rebar tensile strength	Incorrect rebar size changes	Poor surface quality rebars	Incorrect rebar placement and spacing	Incorrect cover to reinforcement	Inadequate overlap distance	Missing reinforcement bars	Loose ties	Un-staggered splices and connections	Deformed or damaged formwork boards	Wrongly dimensioned formwork.	Inadequately sealed formwork joints (not leak proof or /water tight).	Poor formwork board surface quality.	Incorrect formwork elevation.	Insufficient formwork support and bracing	Using inappropriate poker head for vibrating	Erratic concrete delivery to site (cold joint formation)	Inadequate poking/vibration	Over poking	Poor concrete workability	Over workability of concrete	Grout loss	Formwork displacement during placement	Placement under adverse weather conditions	Poor and inadequate curing	Damage during stripping
1	2	3	3	3	4	4	4	4	4	4	2	5	5	3	5	8	6	7	2	6	7	8	7	7	4	6
2	10	8	10	10	10	10	10	10	10	6	10	6	6	5	10	10	5	7	2	8	6	6	6	5	7	7
3	5	6	2	3	3	4	7	8	6	4	5	8	8	7	7	5	4	7	4	7	3	4	8	9	3	4
4	8	8	1	3	9	9	10	1	2	10	10	10	8	2	9	2	2	5	1	9	9	5	1	6	2	2
5	10	9	10	8	10	7	10	7	10	10	8	10	7	10	9	9	10	10	10	10	10	10	10	10	10	5
6	2	4	5	7	8	8	8	8	3	9	9	8	9	5	9	8	6	9	9	8	9	9	8	6	9	2
7	10	9	7	9	8	10	10	8	7	7	7	8	6	6	7	6	7	9	7	7	8	9	8	7	9	8
8	1	2	8	4	8	5	7	4	2	8	2	7	8	2	8	7	4	8	0	8	8	4	7	8	8	2
9	10	10	6	6	5	7	7	5	5	4	8	4	3	10	10	4	6	6	5	8	4	6	4	10	5	5
10	2	3	1	5	7	3	4	1	0	8	9	6	6	6	6	0	1	7	5	0	3	8	9	8	3	2
11	4	6	7	8	7	8	8	8	9	9	9	7	8	9	10	5	8	9	8	9	9	7	10	10	7	5
12	9	8	9	9	9	9	7	6	5	6	5	9	8	4	9	3	7	7	9	8	8	9	9	9	9	7
13	10	9	9	9	9	9	10	9	9	9	10	10	8	6	10	9	10	10	10	10	10	10	10	10	10	8
14	8	9	9	10	5	10	10	8	8	6	10	7	3	10	10	6	8	7	7	6	6	7	10	10	7	2
15	10	10	8	9	9	9	9	7	8	10	10	10	9	10	10	8	10	10	9	8	8	9	10	8	8	7
16	7	4	4	4	8	4	7	6	5	10	9	9	8	10	10	9	7	7	7	5	5	4	8	6	6	4
17	10	10	4	10	8	10	10	5	4	5	7	3	2	8	5	6	3	10	2	7	4	6	4	6	10	6
18	7	8	7	9	8	9	7	8	8	7	9	9	10	8	9	9	9	8	8	9	9	9	7	8	7	9
19	9	7	5	7	6	8	5	8	7	7	6	8	4	5	10	7	7	6	3	8	7	7	8	6	7	10
20	5	10	5	10	9	9	10	9	8	8	10	5	5	10	10	5	6	9	3	6	3	3	6	8	7	5
21	1	3	5	4	5	3	3	1	2	2	2	7	2	7	7	7	6	7	7	5	5	7	8	5	0	0
22	1	2	6	6	6	4	1	2	2	4	1	1	4	1	1	0	6	6	6	6	6	2	1	8	4	2
23	7	7	6	8	8	8	9	7	9	8	8	9	6	7	9	8	9	8	8	6	6	9	8	8	6	6

5.5.1 Fictitious dataset development

As discussed in the preceding chapters, this study aims to develop a prediction model that can promote data-driven decision-making during quality checks on construction sites. To develop the model, various steps were taken to clean and prepare the initial dataset (Table 5.7) to make it suitable for modelling. The following section describes the steps taken to compile the final dataset that was used for modelling starting with the definition of a few key terms.

5.5.1.1 Definition of terms

1. **Continuous variables** – are numeric variables that can have an infinite number of values between any two values.
2. **Discrete variables** - are variables that can have a countable number of values between any two values.
3. **Outliers** – are values/data points that lie at an abnormal distance from other values/data points in a dataset.
4. **Fictitious data** – data simulated from actual events and used for training ML models.

The first step taken to clean the dataset was to investigate whether the development of the model required the 26 identified variables. From the three concreting stages proposed in Chapter 4 of the study (preplacement, placement, and post placement) the variables from the post placement stage were eliminated from the final dataset. This was done as the model was limited to construction activities leading up to concrete placement such as reinforcement works, formwork installation, and the concrete placement processes. The post placement triggers which were excluded from the final dataset were curing and stripping of formwork.

The remaining 24 variables were analysed with some variables grouped under one variable. The grouped variables included some triggers identified during the concrete placement phase. For instance, "over vibrating" and "under vibrating" were grouped under the "vibration consistency" variable. Additionally, "over workability of concrete" and "poor concrete workability" were placed under "concrete workability". After cleaning, the dataset consisted of a total of 18 variables as summarized in Figure 5.6.

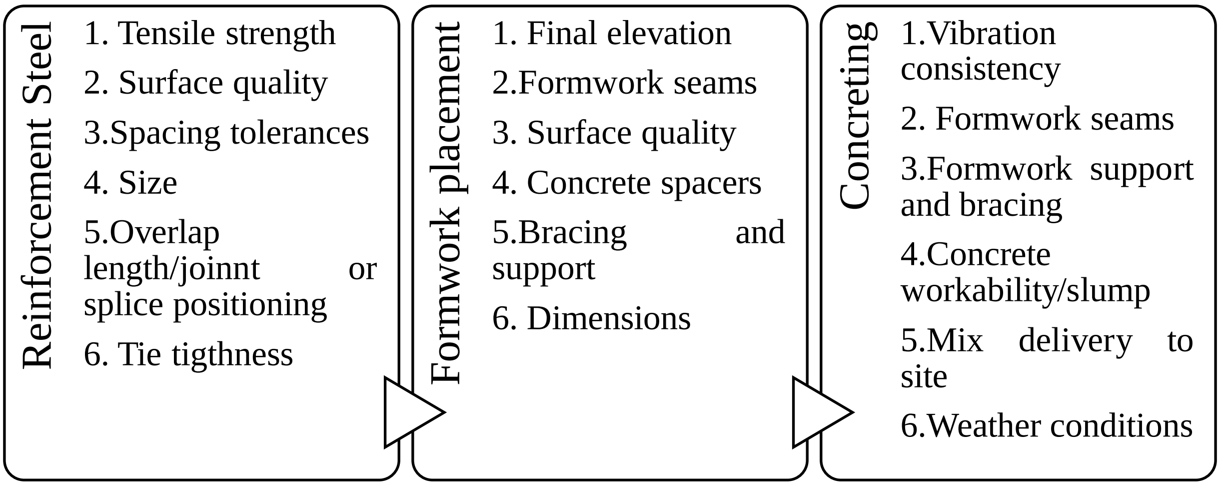

Figure 5.6; Final variables to be considered in the fictitious dataset after processing of data was complete.

After defining the 18 variables to be used in the modelling dataset, a fictitious dataset comprising 1000 random instances was generated in Microsoft Excel. The 10 Likert scale used in the questionnaire survey was also adopted for ratings as illustrated in Table 5.8 with reinforcement steel, formwork, and concreting variables colour coded in yellow, green, and blue respectively. This approach was used for illustrating how a predictive model can be developed.

Table 5.8; Extract of Fictitious dataset with the final 18 variables colour coded by construction stage.

Tensile strength	Surface quality	Spacing tolerances	Size	Overlap distance/ joint or splice	Tie tightness	Final elevation	Seam sealing	Surface quality	Cover to reinforce ment	Bracing and support	Dimensions	Vibration consistency	Formwork support and bracing	Formwork seam tightness	Concrete workability	Mix delivery to site	Weather conditions
9	6	10	3	3	7	1	4	3	6	7	4	8	4	3	8	1	3
3	5	3	6	9	5	4	1	4	10	9	8	9	6	7	3	7	8
4	5	10	7	10	3	3	4	2	7	10	4	4	2	5	9	9	10
2	3	1	1	10	1	2	5	3	1	5	9	8	4	1	1	7	2
4	1	2	1	7	10	7	10	7	6	5	3	8	7	4	4	7	2
1	7	9	5	5	1	8	6	4	10	6	10	7	3	6	4	1	8
3	9	3	10	5	1	9	9	2	7	8	8	4	8	7	6	3	6
4	4	2	4	7	5	4	9	1	4	3	8	7	8	7	10	1	4
5	7	9	2	9	7	5	4	1	1	3	5	7	7	1	2	5	9
3	4	10	4	8	3	8	7	7	6	2	7	2	6	8	4	10	4
6	7	4	4	3	9	9	10	3	8	8	4	6	10	1	2	3	4
10	3	9	9	1	4	7	3	2	8	6	3	5	2	2	9	6	5
4	2	10	8	10	6	9	8	7	6	8	4	9	2	4	8	1	9
7	8	5	6	3	8	5	3	1	9	5	4	5	9	8	8	1	9
2	8	6	9	3	5	6	6	3	5	8	1	3	2	10	1	8	6
10	7	4	9	5	8	1	5	4	3	6	3	9	7	5	9	9	7
4	8	3	6	2	10	1	8	3	2	6	9	6	3	8	7	8	9
4	1	7	5	9	10	10	4	7	6	6	10	6	8	8	5	9	1
10	10	1	9	4	8	5	9	1	7	9	2	2	2	5	10	5	8
8	5	5	7	6	6	6	5	10	6	5	10	6	7	4	5	5	8
7	10	5	4	7	1	8	2	2	4	7	9	10	8	6	5	2	2
6	10	10	10	2	2	1	4	6	6	3	10	8	6	5	8	6	2
7	6	10	2	4	2	9	7	5	7	5	10	5	2	5	5	8	7
4	6	2	3	6	9	6	7	9	5	3	7	6	3	4	6	8	1

Following the proposed categorization of the concrete construction process into three stages, in order to proceed to formwork installation, the reinforcement steelworks should conform to the stipulated standards to minimize the probability of rework occurrence. Taking the reinforcement stage for example, to determine whether to proceed to formwork installation, the column averages from the initial dataset (Table 5.7) were calculated and used as threshold values in modelling. Illustrated in Table 5.9 are the calculated column/variable averages obtained from Table 5.9. For instance, the average rating for the influence of poor rebar surface quality was 6 therefore a rating of 6 and above has high probabilities of resulting in rework. These calculated averages are later used in the development of the predictive model decision matrix as discussed under section 6.2 in the following chapter.

Table 5.9; Column averages for variables in the reinforcement steel stage obtained from Table 5.7 highlighted in blue.

Respondant	Tensile strength	Size	Surface quality	Spacing tolerances	Overlap distance/joints or splices	Tie tighness
1	2	3	3	3	4	4
2	10	8	10	10	10	10
3	5	6	2	3	4	8
4	8	8	1	3	9	1
5	10	9	10	8	7	7
6	2	4	5	7	8	8
7	10	9	7	9	10	8
8	1	2	8	4	5	4
9	10	10	6	6	7	5
10	2	3	1	5	3	1
11	4	6	7	8	8	8
12	9	8	9	9	9	6
13	10	9	9	9	9	9
14	8	9	9	10	10	8
15	10	10	8	9	9	7
16	7	4	4	4	4	6
17	10	10	4	10	10	5
18	7	8	7	9	9	8
19	9	7	5	7	8	8
20	5	10	5	10	9	9
21	1	3	5	4	3	1
22	1	2	6	6	4	2
23	7	7	6	8	8	7
24	6	8	2	4	4	3
25	10	8	5	5	8	5
26	10	10	6	10	7	6
27	10	10	6	10	10	7
28	9	10	8	10	5	5
29	10	7	7	9	6	4
30	10	10	8	7	10	8
31	9	5	9	10	7	8
32	7	8	7	9	7	7
33	8	8	3	10	10	7
	7.18181818	7.24242424	6	7.42424242	7.303030303	6.06060606
Average	7	7	6	7	7	6

The outcome after rating each variable was based on the logic that if any variable input is greater than or equal to the variable average, then it is likely that rework will occur therefore works should not proceed to the next construction stage before amending/rectifying the attributes that do not conform to the expected standard. Additionally, the variables were assigned normalised weights in each category. These weights were assigned on the assumption that the various attributes influence the probability of rework to different extents based on the survey results.

To calculate the weight of each attribute on rework occurrence, the summation of the ratings of the six attributes (reinforcement stage) from the 33 respondents was calculated. To determine the weight of the "*Tensile Strength*" attribute, in particular, the total ratings of that attribute were divided by the total ratings of all six attributes (reinforcement stage) as illustrated in the following equation.

$$Weight\ of\ "Tensile\ Strength"\ attribute = \frac{\sum(Tensile\ strength\ ratings)}{\sum(The\ six\ reinforcement\ attributes\ ratings)}$$

$$Weight\ of\ "Tensile\ Strength"\ attribute\ = \frac{237}{1360}$$

$$= 0.174$$

The weights of the remaining attributes were calculated using the same equation and are summarized in Table 5.10.

Table 5.10; Calculated weights for the 18 variables in the dataset used for the formulation of the predictive model.

Reinforcement Steel	Weight	Formwork	Weight	Concreting	Weight
Spacing tolerances	0.180	Bracing and support	0.185	Vibration consistency	0.320
Overlap distance/ joint or splice position	0.177	Dimensions	0.178	Concrete workability	0.220
Tensile strength	0.174	Seam sealing	0.172	Formwork support and bracing	0.123
Surface quality	0.146	Cover to reinforcement	0.164	Weather conditions	0.122
Size	0.176	Final elevation	0.162	Formwork seam tightness	0.108
Tie tightness	0.147	Surface quality	0.139	Mix delivery to site	0.107
	∑=1[1]		∑=1		∑=1

As mentioned in the sections above, these criteria, based on the feedback from the questionnaire survey were used to develop and test the model discussed in the next chapter. In a real-life project setup, sufficient rework data from actual rework incidences is captured and used for modelling and testing.

[1] Assigned normalised weights.

5.6 Chapter Summary

Descriptive analysis of the data collected was presented in statistical summaries, tables, and graphs, and the most critical risk triggers per construction stage were identified. The flow of stages from data collection to analysis to the presentation of the final dataset discussed in this chapter is summarized in Figure 5.7.

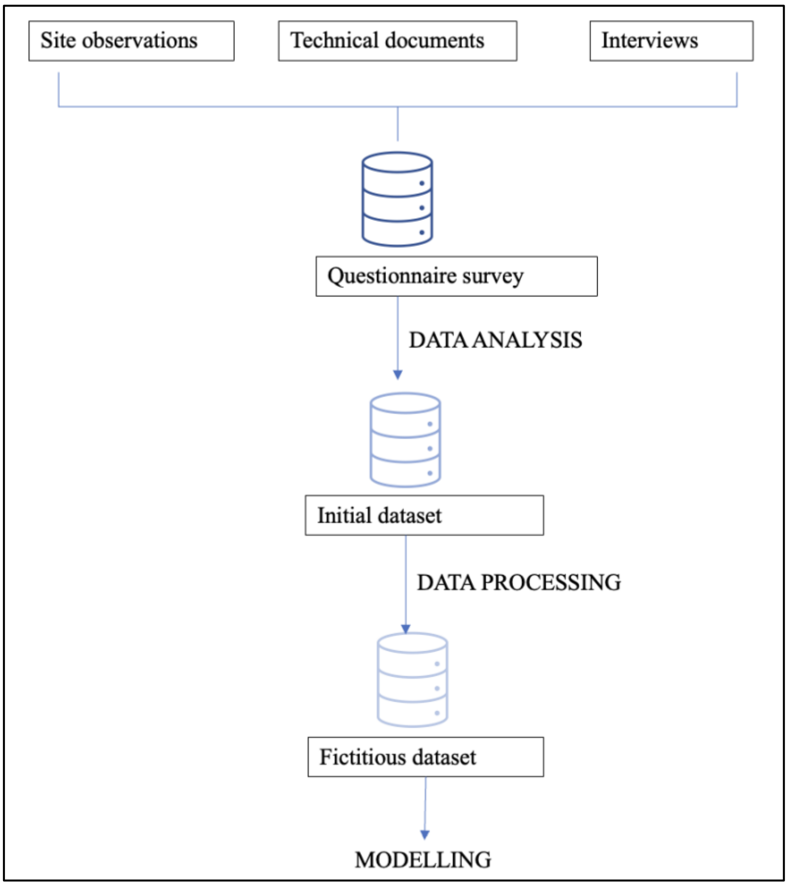

Figure 5.7; Summary of steps and processes discussed in Chapter 5.

The analysis of the survey results revealed that during the placement of reinforcement steel, missing rebars greatly influence structural concrete rework. It was also found that during formwork installation, insufficient support/bracing is the greatest rework trigger which corresponds with findings from literature. Inadequate vibration and formwork displacement during concrete placement were found to also influence the probability of rework occurrence. The feedback ratings from the respondents were captured and presented in Microsoft Excel and specifically used to generate fictitious rework instances for the test model. The detailed processing of the initial dataset discussed in this chapter made use of the tools and accessories introduced in Chapter 3 with the final dataset consisting of 18 variables and 1 000 random instances that will be used for predictive modelling in the following Chapter.

CHAPTER SIX

MODELLING, EVALUATION AND DEPLOYMENT.

6.1 Chapter Overview

This chapter fulfils the aim of the study by providing a step-by-step description of how the fictitious dataset compiled in Chapter 5 was used in the development of a predictive model. The modelling and evaluation of this dataset on five classification algorithms following some steps of the CRISP-DM methodology is also discussed in this chapter. Section 6.4, the deployment phase, describes the development of a predictive data analytics user interface (test model) based on the fictitious dataset and best performing algorithm. The tools and accessories introduced in Chapter 3 are extensively used in this chapter for the development of a model that can be used by organisations to proactively manage structural concrete rework. Due to the unavailability of sufficient project data, hypothetical scenarios are applied in the test model to illustrate how it can aid in decision-making during structural concrete inspections to proactively manage rework occurrence on site. Additionally, the confusion matrix performance test is used to validate the prediction accuracies of the test model on various datasets. Figure 6.1 summaries how the predictive model was developed from the processes described above.

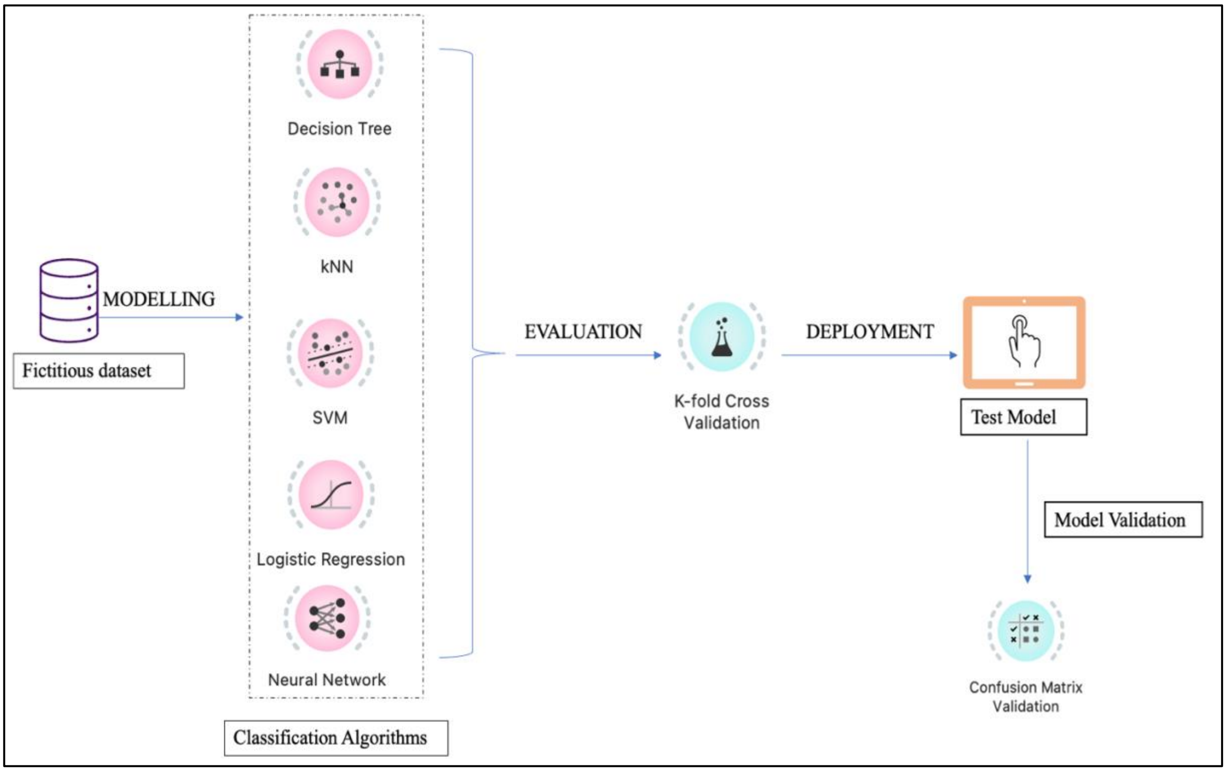

Figure 6.1; Summary of steps and processes taken to develop the predictive model.

6.2 Modelling

Modelling is considered one of the fundamental steps in the CRISP DM methodology as it recognises patterns between data points in a dataset that aid in the development of predictive models. The aim of the study is to apply data analytics in the development of a structural concrete rework prediction model. Therefore, modelling in this study refers to finding a function that maps the identified rework risk triggers to the most appropriate rework outcome at each construction stage with acceptable accuracy and precision. Five classification algorithms were selected due to their maturity and efficiency in the classification of non-linear data science problems as discussed in section 2.7.1.1. Emphasis on the development of the predictive model will be on practicality, efficiency, and reliability.

6.2.1 Design considerations

The 18 input variables from the fictitious dataset (Table 5.8) developed in the previous chapter were converted to continuous numeric variables so that they can be used in the Orange environment for modelling.

The output/outcome variable (decision to "*Proceed to next stage*" or "*Expect rework*") was made a discrete variable so that it can be modelled in classification predictions. The following steps were taken to further clean the dataset for modelling purposes:

1. Data was checked for outliers (any value above 10) and negative numbers.
2. Data was checked to ensure that only numerical values were entered as inputs.

It should be noted that the 18 variables presented in the fictitious dataset were assigned relative weights based on the questionnaire survey results as summarized in Table 5.10. These weights were assigned on the assumption that the various attributes influence the probability of rework to different extents based on the survey results.

After assigning weights to each individual attribute, the decision/outcome "*Proceed to next stage*" was determined by comparing the summation of the weighted attribute ratings (per construction stage) with the summation of the weighted attribute averages (threshold value). The following equations summarize the theoretical logic behind the decision to "*Expect rework*" or "*Proceed to next stage*".

Decision to "*Proceed to next stage*" is based on:

$\Sigma(Scored\ weighted\ attributes\ per\ stage) < \Sigma\ (Weighted\ attribute\ averages\ per\ stage)$..Equation 1

Decision to "*Expect rework*" is based on:

$\Sigma(Scored\ weighted\ attributes\ per\ stage)\ \geq \Sigma\ (Weighted\ attribute\ averages\ per\ stage)$..Equation 2

The dataset was then imported to the Orange environment from the Microsoft Excel spreadsheet it was developed in. Table 6.1 illustrates the reinforcement steel dataset imported in Orange, where there are 1 000 instances with no missing data, 6 variables (referred to as features in Orange) as well as the two output/target variables which are "*Proceed to next stage*" or "*Expect rework*".

Table 6.1; Extract of the fictitious dataset (reinforcement steel attributes) imported in Orange for modelling.

Data Table

Info

1000 instances (no missing data)
6 features
Target with 2 values
No meta attributes

Variables

Show variable labels (if present)
Visualize numeric values
Color by instance classes

Selection

Select full rows

	Outcome	Tensile strength	Surface quality	Spacing tolerances	Size	Overlap distance/ joint or splice position	Tie tightness
1	Expect rework	1.225	0.62	1.476	0.928	0.875	0.873
2	Expect rework	1.75	0.93	1.476	0.812	0.525	0.873
3	Proceed to next stage	0.7	1.24	2.460	0.928	1.4	0.485
4	Expect rework	1.225	0.155	0.492	0.232	0.7	0.194
5	Proceed to next stage	1.75	1.24	2.460	0.348	0.875	0.388
6	Proceed to next stage	1.575	1.55	0.738	0.464	1.225	0.388
7	Proceed to next stage	0.175	1.085	0.738	0.696	0.35	0.485
8	Expect rework	0.7	0.465	1.230	0.812	0.7	0.97
9	Proceed to next stage	0.35	0.93	0.246	0.464	1.225	0.097
10	Proceed to next stage	1.4	1.395	2.460	0.58	0.7	0.291
11	Proceed to next stage	1.4	0.465	2.214	1.044	1.575	0.776
12	Proceed to next stage	0.175	1.085	0.738	0.464	0.7	0.873
13	Proceed to next stage	0.875	0.465	1.230	0.696	0.35	0.291
14	Proceed to next stage	1.575	0.465	1.230	1.044	0.35	0.582
15	Proceed to next stage	1.225	1.55	1.722	1.16	0.525	0.582
16	Proceed to next stage	1.75	1.085	2.214	0.464	0.35	0.679
17	Proceed to next stage	0.35	1.55	0.984	0.928	0.875	0.582
18	Proceed to next stage	0.175	1.24	1.968	0.812	1.4	0.097
19	Expect rework	0.525	1.395	0.492	0.232	0.175	0.776
20	Proceed to next stage	0.35	0.465	0.984	0.812	0.175	0.776
21	Proceed to next stage	1.4	0.93	0.738	1.16	1.05	0.485
22	Proceed to next stage	0.175	0.775	0.246	0.464	0.175	0.776
23	Proceed to next stage	1.75	1.55	1.968	0.58	0.35	0.97
24	Proceed to next stage	1.225	0.465	0.738	0.116	1.05	0.776
25	Proceed to next stage	0.175	0.62	0.246	0.232	1.4	0.097

Modelled using five algorithms as illustrated in Figure 6.2, the classifiers performed predictive calculations on the reinforcement steel training dataset. The data was mapped to the kNN, SVM, Decision Tree, Neural Network, and Logistic Regression classifiers. The kNN model makes predictions according to the nearest training instance whilst the SVM maps inputs to higher dimensional spaces. The Decision Tree and Logistic Regression classifiers predict with forward pruning and ridge regularization respectively. Neural Networks learn and recognize underlying patterns and relationships between data points through a network of nodes. Refer to section 2.7.1 for detailed descriptions of each classification algorithm.

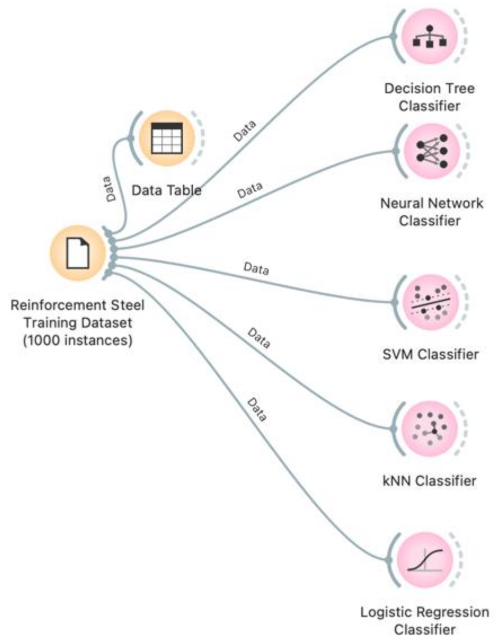

Figure 6.2; Orange canvas showing the mapping of the reinforcement steel dataset to the five classification algorithms.

6.3 Evaluation

After the training dataset was mapped to the five classifiers, the performance of each classification algorithm was evaluated using k-fold cross validation as illustrated in Figure 6.3.

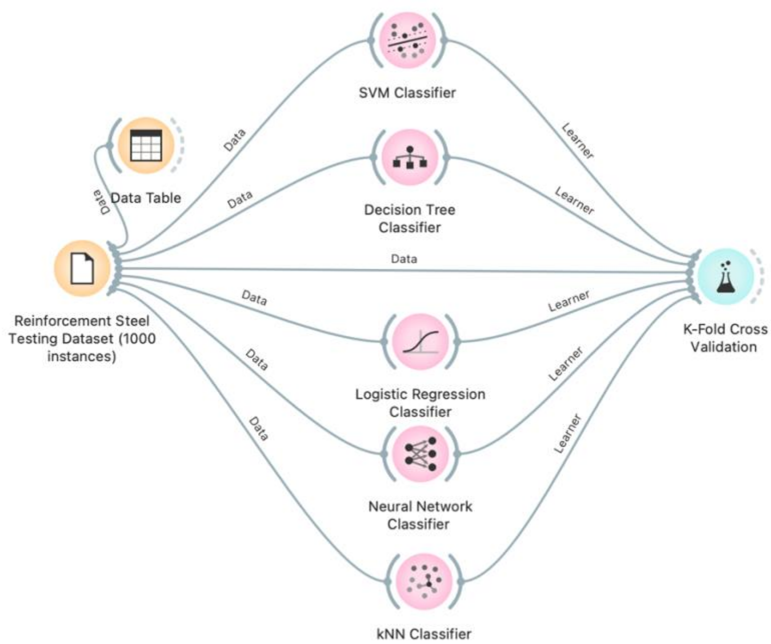

Figure 6.3; Orange canvas showing the mapping of the classifiers to the cross-validation performance test.

The 20-fold cross validation, a statistical method used to estimate the performance of algorithms on new data was used in this study to evaluate the performance of the five classification algorithms. The data sample was randomly split into 20 non-overlapping folds/groups of approximately equal size. Each group/fold was used as a testing set with the remaining 19 groups as training sets. This procedure was repeated and carried out randomly on all the 20 groups/folds. The cross-validation performance test calculated for each classifier the following parameters (described in the following section) as summarized in Table 6.2:

- Area under the ROC curve (AUC) where ROC – Receiver Operating Characteristics indicate the performance of a model in distinguishing the given classes, in terms of the predicted probability.
- Classification accuracy (CA)
- Precision
- Recall
- F1 score

Table 6.2; Orange canvas illustrating the performance of the five classification algorithms after 20-fold cross validation.

Sampling type: Stratified 20-fold Cross validation
Target class: Average over classes

Scores

Model	AUC	CA	F1	Precision	Recall
kNN	0.9106299527735532	0.983	0.9816528480206315	0.9819798647854203	0.983
Tree	0.7827450038688333	0.976	0.9742974059188892	0.9737612841365105	0.976
SVM	0.8922863470209984	0.959	0.9408871873404799	0.9234458917835672	0.959
Neural Network	0.9915152485391819	0.985	0.9838113364887925	0.9843614050558495	0.985
Logistic Regression	0.9306278182448838	0.961	0.9418878123406426	0.9235209999999999	0.961

***Note:** Decimal places could not be adjusted in the Orange Software.

The AUC score measures the level of accuracy in a model with a rating of 1 representing predictions that are 100% correct. The Neural network model recorded the highest AUC score of over 0.99 which is considered high accuracy according to (Dibble, 2020). Both the Logistic Regression and kNN classifiers recorded an AUC rating of about 0.90 with the SVM and Decision Tree classifiers performing with AUC scores of approximately 0.9 and 0.8 respectively.

The Classification Accuracy (CA) score is referred to as the fraction of correctly predicted outcomes with a score of 1 indicating a 100% prediction accuracy. The Neural network algorithm records the highest CA of 0.985. The Neural network classifier also recorded the highest F1 rating which is a weighted mean score of a model's precision and recall (sensitivity). The Precision score indicates the proportion of true positives in instances classified as positive. Recall on the other hand is the proportion of true positives among all positive instances hence high-performance models typically have high recall ratings. The Neural network model recorded relatively high Precision and Recall scores relative to the other models. It can therefore be concluded from the performance indicator results that the Neural network is the best performing model.

6.4 Deployment

Deployment, which is the final process in the CRISP-DM framework involves the application of the developed model for predicting outcomes in new datasets. Deployment can range from compiling a report write up to developing an industry wide model depending on the aim and objectives of the data science project. If the objectives of the data science problem are to increase knowledge, this gained knowledge is to be clearly presented.

In this study, model deployment was executed in Microsoft Excel using Visual Basic for Applications (VBA) programming to develop a user interface that can be used by construction professionals during structural concrete inspections on site to proactively manage rework. The fictitious dataset developed for modelling in Orange was used in the formulation of the user interface with the best performing algorithm (Neural network) running in the test model background for computations.

6.4.1 Prediction (test) model

Due to limited data, the model could not be tested on an actual construction project. Therefore, a hypothetical inspection scenario was be carried out to illustrate how the model can promote data-driven decision making on site. The test model comprises three steps, which integrate input from the user with computations ran in the background of the model by the Neural network algorithm as summarized in Table 6.3.

Table 6.3; Overview of steps of the test model.

Step	Name of step	Summary of step	User input/model background
1	Attribute ratings	User rates conformance of each attribute on a scale of 1 -10.	User input
2	Neural network algorithm maps inputted data	The Neural network algorithm learns the patterns of the input data and maps the most accurate outcome based on patterns learnt from the test data.	Test model background
	Prediction outcome	Model predicts outcome and provides recommendations.	
3	Outcome Summary	Summary of inspection ratings, decision and recommendations made is produced for storage and/or printing.	

6.4.1.1 Step 1

The inspection used in this example is that of reinforcement steel bars of a ground beam. To use the model, it is assumed that the user has sufficient knowledge in structural concrete quality inspections. The first step upon opening the model is a Graphical user interface screen illustrated in Figure 6.4.

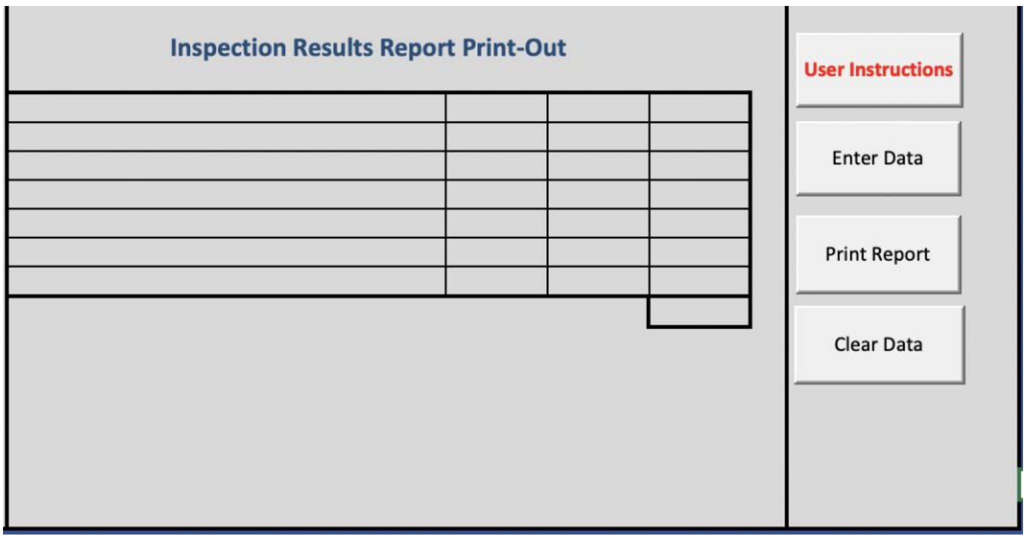

Figure 6.4; Introduction of the graphical user interface used to present the predictive model.

The user selects the "User Instructions" option tab in red text that briefly explains how to use the model. As previously mentioned, this model will evaluate the 18 variables summarized in the fictitious dataset therefore, for this example, the user rates the six attributes under the

"Reinforcement steel" inspection tab. Figure 6.5 illustrates the screen that appears after selecting the "User Instructions" tab.

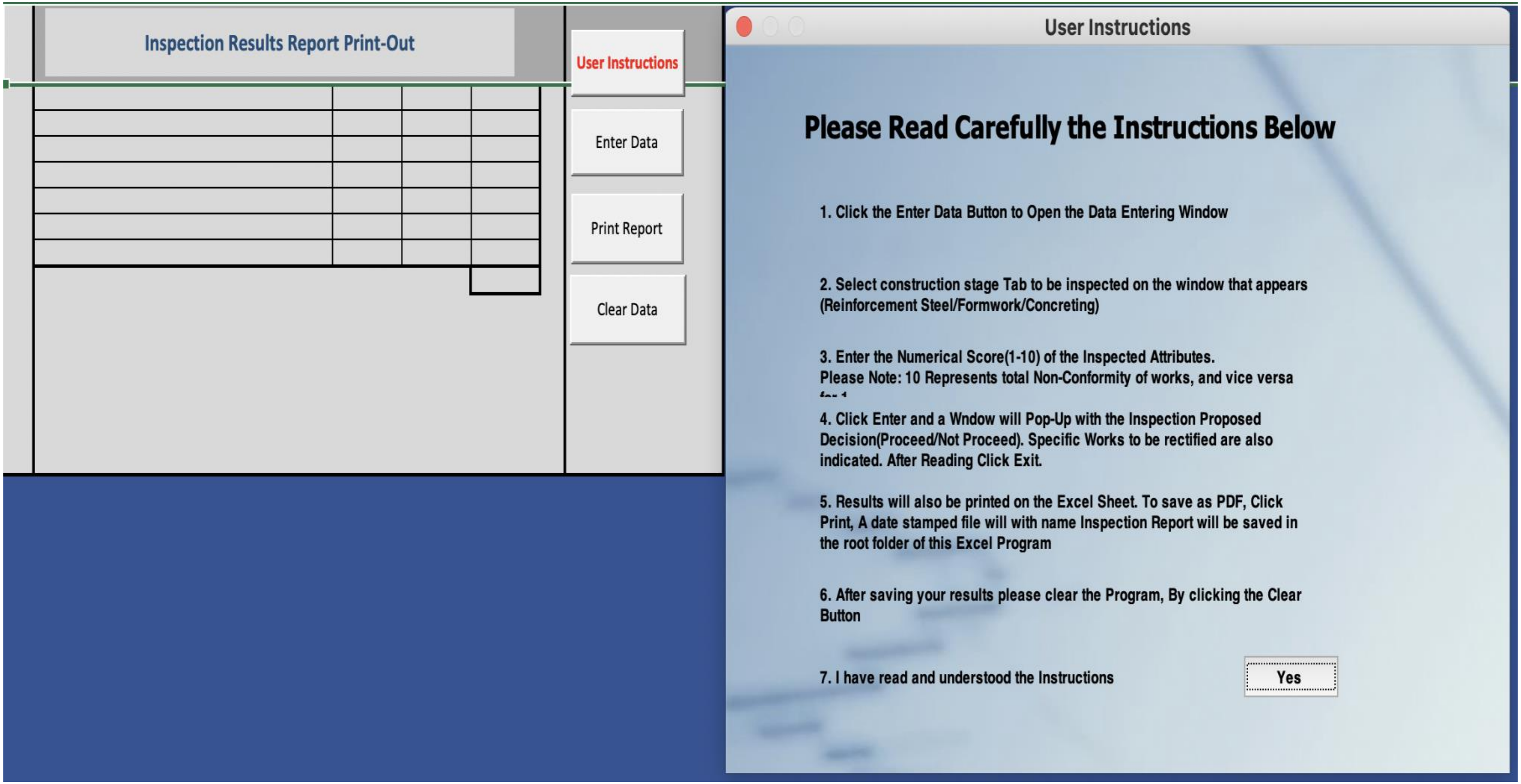

Figure 6.5; Summary of user instructions for the predictive model.

The “Reinforcement Steel” tab contains 6 variables that a user is required to rate as illustrated in Figure 6.6. In this example, rebars of a larger size were used instead of the ones specified in the design drawing. The user then rates an 8 since it is close to a non-conformance. The tensile strength and rebar spacing are good hence the ratings of 1. The surface quality, position of mechanical splices, and tie tightness are satisfactory hence the scores of 5.

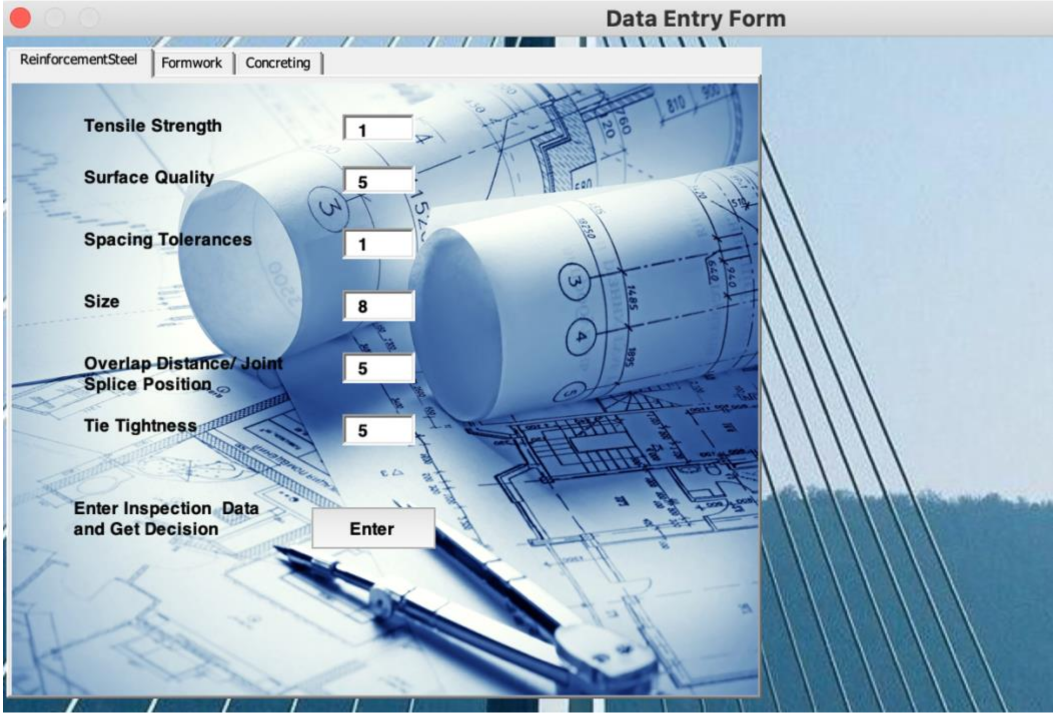

Figure 6.6; Illustration of how a user rates the selected variables during structural concrete reinforcement steel inspections.

6.4.1.2 Step 2

After rating the attributes, the system captures the scores and multiples them with the assigned weights. The weights ensure that the model unevenly allocates the influence of each of the attributes on rework probability. The weighted scores are mapped on the Neural network algorithm which predicts an outcome and provides recommendations based on the pattern learnt from the testing dataset. Figure 6.7 illustrates the decision generated and this example suggests that the user rectifies the rebar sizes although a low chance of rework is predicted (as other variables are satisfactory). Since the rating of the reinforcement attribute was close to a non-conformance (8) and greater than the average rating of 7(for reinforcement bar sizes from table 5.9), the model recommends rectifying this attribute.

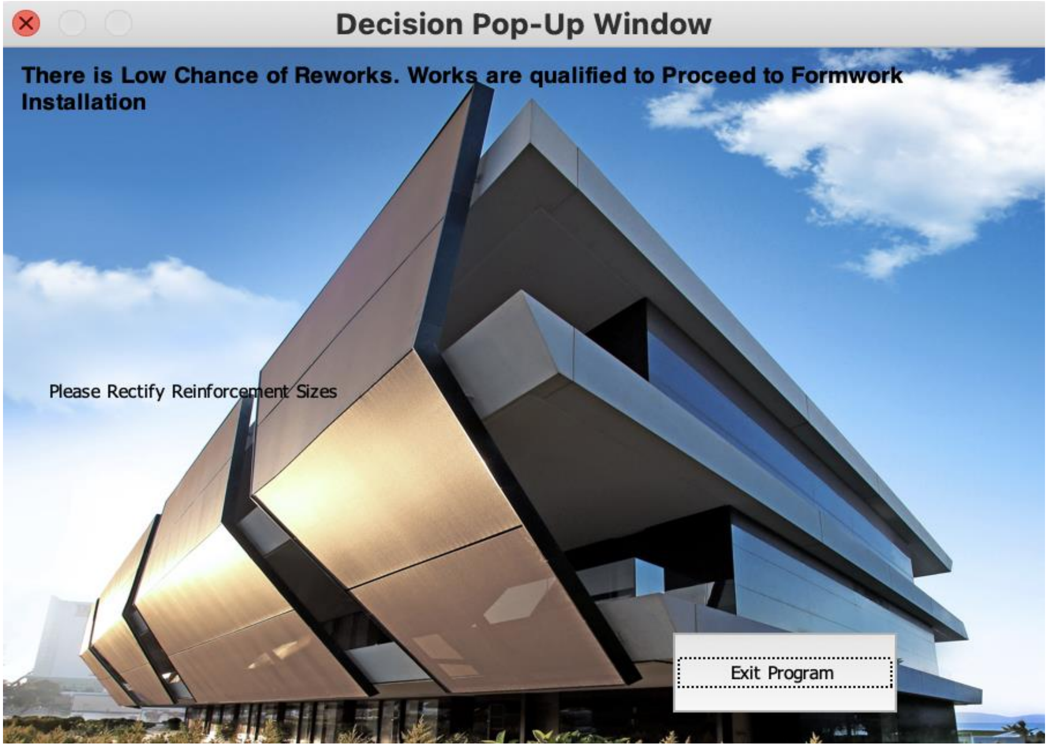

Figure 6.7; Predicted decision and recommendations after user rates all the variables.

6.4.1.3 Step 3

After the model predicts an outcome, the inspection ratings are presented in a report that can be printed or stored in the cloud as illustrated in Figure 6.8.

Inspection Results Report Print-Out

Inspection Atrribute	Attribute Score	Weight	Weighted Attribute
Tensile Strength	1	0.174	0.174
Surface Quality	5	0.147	0.735
Spacing Tolerances	1	0.18	0.18
Size	8	0.176	1.408
Overlap Distance/ Joint Splice Position	5	0.177	0.888
Tie Tightness	5	0.147	0.735

Reinforcement Tensile Strength Passed
Reinforcement Surface Quality Passed
Spacing Tolerances Passed
Re-Check Reinforcement Sizes
Overlap Distance Passed
Tie Tightness Passed

Works Qualified to Proceed to the next stage

User Instructions
Enter Data
Print Report
Clear Data

Figure 6.8; Summary of the reinforcement steel inspection results and recommended decision.

6.4.1.4 Model Validation

To determine the effectiveness and accuracy of the developed model, ideally, real-life rework outcomes are used as inputs in the model and a database based on these entries is compiled. However, due to the lack of access to a running project, validation of the developed model was done with simulated rework datasets. Six datasets of varying size and known outcomes were formulated and fed into the test model to determine the accuracy of the predictions. The outcomes of the simulated datasets were determined using the theoretical equations introduced in section 6.2 which are:

Decision to "*Proceed to next stage*" for validation dataset is based on:

$\Sigma(Scored\ weighted\ attributes\ per\ stage) < \Sigma\ (Weighted\ attribute\ averages\ per\ stage)$..Equation 1

Decision to "*Expect rework*" for validation dataset is based on:

$\Sigma(Scored\ weighted\ attributes\ per\ stage)\ \geq \Sigma\ (Weighted\ attribute\ averages\ per\ stage)$..Equation 2

The confusion matrix was used to evaluate the performance of the developed predictive model on six different datasets. Additionally, this validation test was carried out to obtain a reasonable indication of the dataset size required for accurate predictive modelling. Figure 6.9 illustrates the mapping of the first validation dataset with 100 instances.

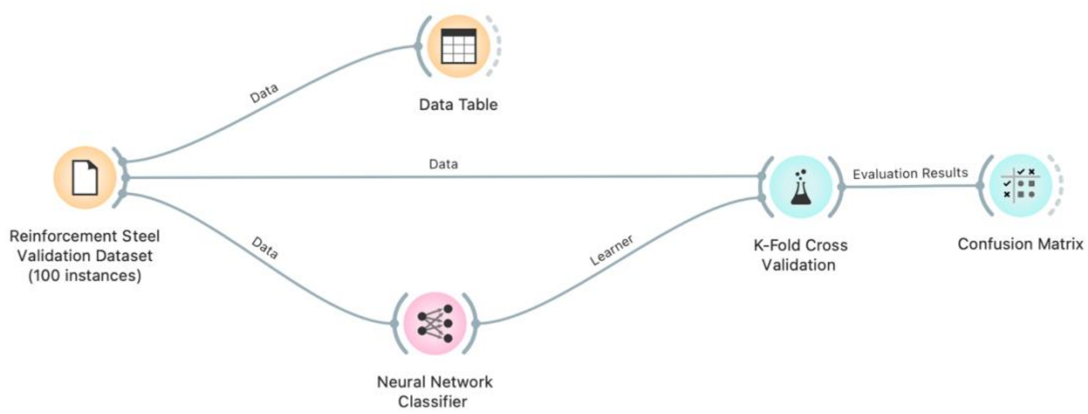

Figure 6.9; Section of the Orange canvas showing how the validation dataset is mapped to the confusion matrix performance test widget.

As previously discussed, the dataset has two discrete outcomes per element, which are either to "*Proceed to the next stage*" or "*Expect rework*". For the first validation dataset with 100 instances, 97 of the 100 outcomes were "*Proceed to next stage*" and 3 were "*Expect rework*". After running the dataset through confusion matrix validation, the Neural network model predicted 100 "*Proceed to next stage*" and zero "*Expect rework*" as illustrated in Table 6.4.

Table 6.4; Orange canvas showing the results of the confusion matrix performance test on the validation dataset containing 100 instances.

		Predicted		
		Expect Rework	**Proceed to next stage**	**Σ**
Actual	**Expect Rework**	0	3	**3**
	Proceed to next stage	0	97	**97**
	Σ	**0**	**100**	**100**

To interpret the confusion matrix in Table 6.5 the following terms will be defined.

1) True Positive (TP) – these are instances when the predictive test model correctly predicts the "*Proceed to next stage*" outcome.
2) True Negative (TN) - these are instances when the predictive test model correctly predicts the "*Expect Rework*" outcome.
3) False Positive (FP) – these are instances when the predictive test model predicts "*Proceed to next stage*" when the actual outcome is "*Expect Rework*"
4) False Negative (FN) – these are instances when the predictive test model predicts "*Expect Rework*" when the actual outcome is "*Proceed to next stage*".

The summation of row elements as we move to the right of the table represents the actual outcomes whilst the column summations indicate outcomes predicted by the test model. Table 6.5 illustrates that for the outcome "*Proceed to next stage*" the model, out of 97 True Positives the test model predicted 100 "*Proceed to next stage*" outcomes. This implies the test model misclassified 3 outcomes as False Positives that is, 3 outcomes were predicted as "*Proceed to next stage*" which in actual fact were "*Expect Rework*". The model failed to classify the 3 "*Expect Rework*" outcomes hence it has zero True and False Negatives.

The misclassified predictions (False Positives and False Negatives) can be used to calculate the misclassification/error rate of the prediction model. This error rate is an indication of the prediction accuracy of the model, that is, a low error rate translates to a model of high accuracy. For the validation dataset in 100 instances, the error rate is calculated as;

$$Misclassification\ rate = \frac{False\ Positives\ (FP) + False\ Negatives\ (FN)}{Total\ number\ of\ outcomes} \text{..Equation 3}$$

$$Misclassification\ rate\ (100\ instances) = \frac{3+0}{100} = 0.03$$

This error/misclassification rate also indicates how often the algorithm makes false predictions, which in this case is 3 times in every 100 instances. The confusion matrix validation was then run on five more datasets with 500, 1000, 2500, 5000, and 10000 randomly generated instances with the results for each dataset summarised in Table 6.5. The confusion matrix results for other datasets are attached in Appendix H of this study.

Table 6.5; Summary of confusion matrix results for five different datasets.

Size of dataset		**Proceed to next stage outcome**		**Expect rework outcome**		**Error rate**
		Actual	Predicted	Actual	Predicted	
1	100	97	100	3	0	0.0300
2	500	490	498	10	2	0.0200
3	1 000	969	981	31	19	0.0140
4	2 500	2 409	2 416	91	84	0.0140
5	5 000	4 835	4 855	165	145	0.0076
6	10 000	9 655	9 664	345	336	0.0009

The results from the confusion matrix on the six datasets illustrate that the overall performance of the prediction model is influenced by the size of the dataset. Although the algorithm is not 100% accurate as illustrated by the number of misclassified outcomes the error/miscalculation rate decreases as the dataset increases. Figure 6.10 illustrates that larger datasets have smaller misclassification errors, therefore, making them more accurate.

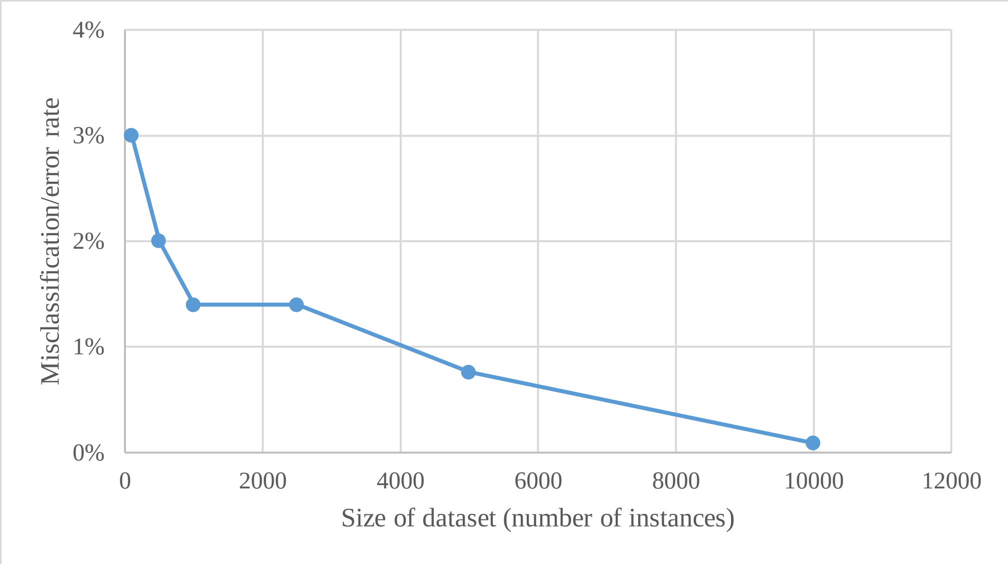

Figure 6.10; Illustration of how the misclassification rates vary with dataset size (Computed from the six validation datasets confusion matrix test results).

6.4.2 Proposed industrial model

The test model described in the previous section is based on a fictitious dataset compiled from the findings of a survey conducted on structural concrete experts. Additionally, the decision outcome "*Proceed to next stage*" or "*Expect rework*" is based on a theoretical threshold value. Due to insufficient historical project data, this fictitious dataset was used to illustrate the type of rework data that can be captured by individual organisations in order to develop similar predictive models that can aid in rework management.

For organisations to develop similar models based on actual project data, they should consider effectively capturing rework data on site and processing the dataset to make it suitable for machine learning algorithms. The major difference between the model developed in this study and the one that organisations can use is that the classification algorithms in the envisioned industrial model recognise patterns based on actual project data and not fictitious datasets. This enables the development of robust predictive models based on real life historical data.

Additionally, the industrial model does not require assumptions (weights/predetermined threshold values) as the outcome will be based on the patterns learnt from actual events. Table 6.6 provides a summary of the major differences between the test model developed in this study and an industrial model individual organisations can develop from actual project databases.

Table 6.6; Differences between test model developed in this study and the proposed industrial model.

	Test model developed in this study	**Proposed industrial model**
1	Based on a fictitious dataset formulated from survey results.	Based on actual project data/events.
2	Decision based on patterns learnt by algorithm from fictitious data and predetermined threshold values.	Decision based on patterns learnt by actual rework outcomes.
3	Limited to risk triggers investigated in this study.	Can integrate other rework triggers not considered in this study such as experience of workforce and inputs from other project departments such as resources management.
4	Validated using simulated datasets.	Can be validated with actual project data on an ongoing basis.

6.5 Chapter Summary

This chapter discussed the modelling, evaluation, and deployment techniques employed in Orange and Microsoft Excel in order to accomplish the aim of the study. Five classification algorithms were modelled and evaluated using the 20-fold cross validation performance measure which concluded that the Neural network classification algorithm was the best performing with accuracy and precision of over 95%. After establishing the best performing model, a test model based on the Neural network algorithm was developed to illustrate how it can be implemented in a real-life project scenario.

The test model followed the three steps summarised in Figure 6.11 to predict the outcome of new data inputs. The performance of the developed Neural network model was validated using the confusion matrix performance measure. Six datasets of varying sizes ranging from 100 to 10 000 datapoints were evaluated, and the results indicated that larger datasets produce better performing models with smaller error/ miscalculation rates. Furthermore, results from the confusion matrix test revealed that a dataset of at least 1 000 instances is suitable for accurate predictive modelling. It is therefore recommended to capture as much structured data as possible to develop models of high accuracy.

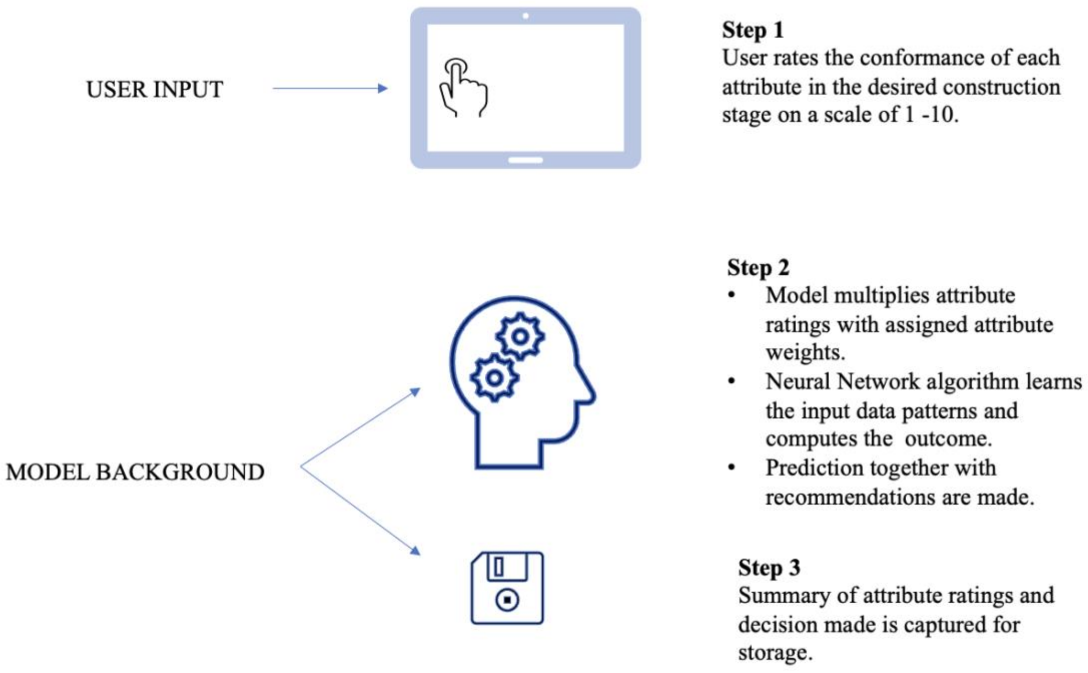

Figure 6.11; Overview of the steps taken in using the test model.

This test model can be used as a guide by individual organisations in the construction industry to effectively capture structural concrete rework data. While existing literature provides methodologies to manage rework based on statistical methods and interproject learning, the model described in this chapter aims to proactively manage rework through the use of classification algorithms bridging the gap between rework management and data analytics. It is not essential for practitioners to understand how classification algorithms work, but rather to understand the data input requirements and possible model limitations.

CHAPTER SEVEN

CONCLUSIONS AND RECOMMENDATIONS

This Chapter outlines the overall summary of the study, the conclusions from the research objectives and proposes areas for further research in line with the implementation of data analytics and data-based decision-making in the construction industry.

7.1 Summary

This research focused on the application of predictive data analytics in the development of a model that can be used as a quality control measure in the management of structural concrete rework on site. Limited to structural elements such as beams, columns, and slabs, the study investigated the risk triggers that potentially result in structural concrete rework in a typical construction project. The case study, Hwange Expansion Project, a mega thermal power plant construction site was the primary source of data. Spanning a period of twelve months, a mixed data collection approach was adopted which included weekly site inspections by the researcher, reviews of site technical documentation as well as expert interviews with concrete construction professionals at the Hwange Expansion Project construction site.

Data collected from the above-mentioned methods was used to formulate a questionnaire survey aimed at professionals with experience in structural concrete construction. The findings from the survey were used to compile a fictitious dataset that was used for modelling, due to the lack of sufficient historical project data to develop a predictive model. Criteria to predict the likelihood of rework occurrence based on the fictitious dataset was also established. The final model consisted of 18 attributes from three structural concrete construction phases namely:

1. Reinforcement steel placement
2. Formwork installation
3. Concrete placement

Before modelling, various prepossessing techniques were applied to the fictitious dataset to make it compatible with the Orange software adopted due to its vast collection of data visualisation tools and algorithms for data mining and analysis. The data was trained on five classification algorithms and the 20-fold cross validation performance test was used to evaluate the performance of each classifier. The 20-fold cross validation performance test split the sample into 20 folds dividing the dataset into 20 subsets used for training and testing the classifiers. The Neural network classifier was found to be the best performing model with an accuracy and precision of over 95%.

The fictitious dataset used for training the classification algorithms was also used to develop a predictive model that can be used to promote data-driven decision-making during structural concrete inspections. The performance of the test model based on the Neural network algorithm was validated by mapping six datasets of varying sizes (from 500 to 10 000 data points) to the confusion matrix test for evaluating the accuracy in predictions of the Neural network test model. The validation results indicated that larger datasets produce better performing models with smaller error/ miscalculation rates.

This predictive model seeks to augment the standard structural concrete quality checks by minimizing human subjectivity in decision-making onsite. The model can also be used as a guide by individual organisations in the construction industry to effectively capture structural concrete rework data. To use the model, it is not essential for practitioners to understand how classification algorithms work, but rather, to understand the data input requirements and possible model limitations. Additionally, this model illustrates that with sufficient historical data, individual organisations can proactively manage rework through similar models.

7.2 Conclusions

Data analytics was successfully applied in the development of a predictive model that can be used for rework management during structural concrete inspections. Motivated by the negative effects rework poses on project cost, schedule, and quality, this study developed a machine learning model that can aid decision-making during structural concrete quality inspections. The study showed that if sufficient historical project data is captured, predictive models can be developed and used to enhance project performance.

The four objectives of the study were achieved as described below:

Objective 1: To identify construction risk triggers that lead to structural concrete rework during the project construction phase.

- This objective was achieved through conducting site observations, technical documents and a literature review, expert interviews as well as the questionnaire survey.
- Some of the risk triggers identified included missing reinforcement steel bars, inadequate formwork bracing and insufficient fresh concrete compaction.
- The results of this objective can serve as a guide to individual organisations about the type of project data to capture to develop similar prediction models.

Objective 2: To categorize the identified risk triggers into phases during reinforced concrete structure construction.

- The second objective was achieved through a thorough review of literature as well as semi-structured interviews with structural concrete experts.
- Categorizing rework risk triggers allowed the easy capturing of data which can also be used by individual organisations as a template to capture historical project data.
- The primary categories identified in the study were:
  1. Pre concrete placement (steel reinforcement works and formwork installation).
  2. Concreting (placement of fresh concrete).
  3. Post concrete placement (curing and stripping of formwork).

Objective 3: To develop a suitable dataset that can be used to train and test the performance of classification algorithms.

- Due to insufficient historical data to develop a predictive model, this objective was achieved by consolidating rework data captured through site observations, reviews of technical documents as well as interviews to compile a questionnaire survey.
- A fictitious dataset based on the findings of the questionnaire survey was formulated, modelled, and trained on five classification algorithms.
- Various assumptions were made on the fictitious dataset such as assigning weights to each variable (based on results from the survey), which, in a real-life project set up would be based on actual project data.
- The model's decision matrix (*which provided an outcome on whether to "proceed to the next construction phase" or "expect rework"*) was based on predetermined threshold values computed from assigned weights and rework trigger attribute averages (from the survey results).

Objective 4: To apply predictive data analytics in developing a model that can be used to reduce the occurrence of construction field rework during construction.

- Various classification algorithms were trained on the fictitious dataset and the Neural Networks model presented the model with the highest classification accuracy and performance.
- Datasets of varying sizes ranging from 100 to 10 000 datapoints were compiled and used for validating the Neural Networks model to assess its prediction accuracy. It was found, as expected, that the accuracy of a model increases with more data points and

that a dataset with at least 1 000 entries would be suitable for accurate predictive modelling.

- The developed model can be used as a quality control management technique that supports data-driven decision-making proactively managing structural concrete rework.

Overall, it can be concluded that capturing sufficient structured project data is essential in the development of robust predictive models in the management of structural concrete rework.

7.3 Research Recommendations.

This section proposes areas for future research both in the construction industry and future studies.

7.3.1 Recommendations for the industry

Implementing data analytics in disciplines with limited available data can be challenging. The construction industry is one such discipline, with limited data required to prototype and incorporate machine learning and artificial intelligence into the day-to-day activities of construction projects. It is therefore beneficial for individual organisations and the construction industry to effectively capture data throughout the project life cycle (initiation to completion). This will ensure sufficient historical data that can be used to develop robust prediction models.

This study demonstrated how with, effective data capturing methodologies, individual organisations can apply data analytics to enhance project performance through predictive models. Construction companies are encouraged to develop systems that can capture and store project data as construction sites are an abundant source of data, spanning from planning, design, and engineering, with over 90% of collected data going unutilized (Snyder et al., 2018).

Project data can effectively be captured through the adoption of Building Information Modelling (BIM) which integrates data across all the phases of the construction project in one central server. Data during project construction can also be collected in real time through the use of various technologies such as laser scanners and concrete curing sensors. This will provide a large database based on real events that provides better insights for machine learning and predictive data analytics. This real time data can be collected from more than one project to enable the development of a diverse dataset that can be used industry wide.

7.3.2 Recommendations for future research

Due to insufficient historical project data, this study identified the common risk triggers that influence structural concrete rework and used these findings to compile a dataset used for modelling. The model's decision matrix (*which provided an outcome on whether to "proceed to the next construction phase" or "expect rework"*) was based on predetermined threshold values computed from assigned weights and rework trigger attribute averages (from the survey results). This criterion was adopted to illustrate the type of data required for predictive data analytics, however, further research can be conducted to establish a decision matrix based on actual project data. As such, individual organisations can develop similar models by mapping historical project data on classification algorithms that recognize patterns in the data. Emphasis should be put in effectively storing and processing the data collected so that it can be suitable for predictive modelling. The introduction of real time data capturing techniques is also recommended as an alternative method for data collection. This can be achieved by using Unmanned Aerial Vehicles (drones) that conduct real time site monitoring and defect detection during the undertaking of the works rather than after inspections at hold points.

The model could not be tested and validated on a real construction site, due to the lack of sufficient data. Validation was then carried out on five fictitious datasets of various sizes ranging from 100 to 10 000 data entries. It is recommended to validate a model with real-life project data as this shows the accuracy of a model in real life cases and not hypothetical scenarios. Future research can also capture data such as the experience of the foreperson and labourers, temperature during the concrete placement as well as the elevation of the cast structural element to use for modelling to develop effective and robust prediction tools that consider other variables that influence structural concrete rework.

REFERENCES

Abdul-Rahman, H., 1995. The cost of non-conformance during a highway project: A case study. *Construction Management and Economics*, 13(1), pp. 23-32.

Ackermann, F., Eden, C. & Williams, T., 1997. Modelling litigation: Mixing qualitative and quantitative approaches. *Interfaces,* 27(2), p. 48–65.

Alwi, S., Hampson, K. & Mohamed, S., 2001. Effect of quality supervision on rework in the Indonesian context. *Asia Pacific Building and Construction Management Journal*, 6(9), pp. 1024-9540.

Arundachawat, R., Roy, R., Al-Ashaab, A. & Shehab, E., 2009. *Design Rework Prediction in Concurrent Design Environment: Current Trends and Future Research Directions,* United Kingdom: Cranfield University .

Ashford, J., 1992. *The management of quality in construction.* London: E & F Spon.

Barber, P. et al., 2000. Quality failure costs in civilengineering projects. *International Journal of Quality and Reliability Management,* 17(4), pp. 479-92.

Basbeth, F., 2014. "Partnering and project performance" The mediating role of innovation and TQM practice in high rise building construction projects , Jakarta, Indonesia. *International Journal of Economics, Commerce and Management,* 4(4).

Battikha, M. G., 2008. Reasoning Mechanism for Construction Nonconformance Root-Cause Analysis. *ASCE Journal of Construction Engineering and Management*, 134(4).

BRE, 1981. *Quality Control on Building Sites*, London. : Building Research Establishment .

BRE, 1982. *Quality in Traditional Housing – An Investigation into Faults and their Avoidance*, Garston: Building Research Establishment.

Burati, J., Farrington, J. & Ledbetter, W., 1992. Causes of Quality Deviations in Design and Construction. *ASCE Journal of Construction Engineering and Management,* Volume 118, pp. 34-49.

Busby, J., 2001. Error and Distributed Cognition in Design. *Design Studies,* Volume 22, pp. 233-254.

Campanella, J. & Corcoran, F., 1983. Principles of Quality Costs. *Quality Progress,* 16(4), pp. 16-22.

Carneiro da Rocha, J. B. & Timóteo de Sousa, R., 2010. Identifying bank frauds using CRISP-DM. *International journal of computer science & information Technology (IJCSIT),* 2(5).

Chan, D. & Kumaraswamy, M., 1997. A comparative study of causes of time overruns in Hong Kong construction projects. *International Journal of Project Management,* 15(1), pp. 55-63.

Cnuddle, M., 1991. Lack of quality in construction – economic losses. *European Symposium on Management, Quality and Economics in Housing and other Building Sectors,* p. 508–15.

Construction Industry Institute, C., 1989. *Costs of quality deviations indesign and construction,* Texas: The University of Texas at Austin, Austin.

Construction Industry Institute, C., 2001. *"The field rework index: Early warning for field rework and cost growth.",* Texas: Rep. NO. RS153-1, Austin.

Construction Industry Institute, C., 2001a. *The field rework index: Early warning for field reworks and cost growth. Research Summary,* Texas: The University of Texas at Austin, Austin.

Construction Industry Institute, C., 2001b. *An investigation of field rework in industrial construction, research summary,* Texas: Construction Industry Institute, Austin, Texas, USA.

Construction Industry Institute, C., 2005. *Making zero rework A reality. RS 203-1 (November),* Texas, The University of Texas at Austin, Austin.

Construction Owners Association of Alberta, C., 2001. *Field Rework Committee meeting minutes,* Alta. Edmonton, Alta.

Construction Owners Association of Alberta, C., 2006. *Project Rework Reduction Tool,* Alberta: Construction Owners Association of Alberta.

Cooper, K., 1980. Naval shipyard production: A claim settled and a framework built. *Interfaces,* 10(6), p. 30–36.

Cooper, K., 1993. The rework cycle: Benchmarking for the project manager. *Project Management Journal,* 24(1), p. 17–22.

Craig, D., 2018. *Most Common Rework Causes and How to Prevent Them.* [Online] Available at: https://www.procore.com/jobsite/construction-rework-causes-costs/#:~:text=According%20to%20%E2%80%9CLearning%20Practices%20as,million%20spent%20on%20a%20job. [Accessed 6 October 2020].

Cusack, D., 1992. Implementation of ISO 9000 in construction. *ISO 9000 Forum Symposium, the International Group for Lean Construction.* Gold Coast, Australia, pp. 157-167.

Data Science Project Management, 2020. *CRISP-DM.* [Online]
Available at: https://www.datascience-pm.com/crisp-dm-2/
[Accessed 22 September 2020 2020].

Davis, K., Ledbetter, W. & Buratti, J., 1989. Measuring design and construction quality costs. *ASCE Journal of Construction Engineering and Management,* 115(3), p. 389–400.

Dibble, M., 2020. *An explanation of machine learning models even you could understand.* [Online] Available at: https://thenextweb.com/neural/2020/04/25/machine-learning-models-explained-to-a-five-year-old-syndication/ [Accessed 19 October 2020].

Eden, C., Williams, T. & Howick, S., 2000. The role of feedback dynamics in disruption and delay on the nature of disruption and delay (D&D)in major projects. *Journal of the Operational Research Society,* 51(3), pp. 291-300.

Egan, J., 1998. *Rethinking Construction – The Report Of The Construction Task Force.* London: Department Of Environment, Transport & The Regions.

Elmousalami, H. H., 2019. Artificial Intelligence and Parametric Construction Cost Estimate Modeling: State-of-the-Art Review. *ASCE Journal of Construction Engineering and Management,* 146(1).

Fayek, A., Dissanayake, M. & OCampero, 2003. *Measuring and classifying construction field rework: A pilot study. Construction Owners Association of Alberta (COAA),* Canada, The University of Alberta, Edmonton.

Feng, P., Tommelein, I. & Booth, L., 2008. *Modeling the effect of rework timing: Case study of a mechanical contractor. In Tzortzopoulos, P. and Kagioglou, M. (Editions 2008).* Manchester, UK, Proceedings of the 16th Annual Conference of the International Group for Lean Construction (IGLC 16), .

Forcada, N., Gondolells, M., Casals, M. & Macarulla, M., 2014a. Factors Affecting Rework Costs in Construction. *ASCE Journal of Construction Engineering and Management,* 143(8).

Gondia, A., Siam, A., El-Dakhakhni, W. & Nassar, A. H., 2019. Machine Learning Algorithms for Construction Projects Delay Risk Prediction. *ASCE Journal of Construction Engineering and Management,* 146(1).

Gunawardena, C. N. et al., 2004. Building an online wisdom community: A transformational design model. *Journal of Computing in Higher Education,* 15(2), p. 40–62.

Hammarlund, Y. & Josephson, P., 1991. Sources of quality failures in building,. *Proceedings of the European Symposium on Management, Quality and Economics in Housing and other Building Sectors,* Lisbon, Portugal, p. 671–9.

Hancock, B., Ockleford, E. & Windridge, K., 2009. An Introduction to Qualitative Research. The NIHR RDS EM / YH.

Helmy, K., 2017. *Inspection, Maintenance and Repair of Structures* , s.l.: Arab Academy for Science, Technology and Maritime Transport.

Hong, H. & He, W., 2015. Effect of Human Error on the Reliability of Roof Panel Under Uplift Wind Pressure. *Structural Safety,* Volume 52, pp. 54-65.

Huovila, P. & Koskela, L., 1997. *"On Foundations of Concurrent Engineering.",* London: Concurrent Engineering in Construction CEC.

Hwang, B. G., Thomas, S. R., Haas, C. T. & Caldas, C. H., 2009. Measuring the impact of rework on construction cost performance. *ASCE Journal of Construction Engineering and Management,* 135(3), pp. 187-198.

IGCSE ICT, n.d. *Expert Systems | IGCSE ICT.* [Online]
Available at: https://www.igcseict.info/theory/7_2/expert/index.html
[Accessed 19 october 2020].

Incze, R., 2019. *The Cost of Machine Learning Projects.* [Online]
Available at: https://medium.com/cognifeed/the-cost-of-machine-learning-projects-7ca3aea03a5c [Accessed 06 October 2020].

ISO.9000, 2005. *Quality management systems — Fundamentals and vocabulary,* Switzerland: International Organization for Standardization.

Josephson, P. & Hammarlund, Y., 1999. The causes and costs of defects in construction. A study of seven building projects. *Automation in Construction,* 8(6), pp. 681-642.

Josephson, P., Larsson, B. & Li, H., 2002. Illustrative benchmarking rework and rework costs in Swedish construction industry. *ASCE Journal of Management in Engineering,* 18(2), pp. 76-83.

Kakitahi, J., Landin, A. & Alinaitwe, H., 2013. An exploratory study of rework causality in Uganda. *Construction Innovation,* 13(3), pp. 266-280.

Koskela, L., 1992. *Application of the New Production Philosophy to Construction.Technical Report.Center for Integrated Facility Engineering.* Department of Civil Engineering., California: Stanford University.

Llozor, D., Okoroh, M. & Egbu, C., 2004. Understanding residential house defects in Australia from the State of Victoria. *Building and Environment* , 39(3), pp. 327-337.

Love, P., 2002a. Influence of project type and procurement method of rework costs in building construction projects. *ASCE Journal of Construction Engineering and Management,* 128(1), pp. 1-29.

Love, P., 2002. Auditing the indirect consequences of rework in construction: Case based approach. *Managing Auditing Journal,* 17(3), pp. 138-146.

Love, P., 2019. Making sense of rework and its unintended consequence in projects: The emergence of uncomfortable knowledge. *International Journal of Project Management,* 37(3), p. 501–516.

Love, P. et al., 2016. The praxis of rework mitigation in construction. *ASCE Journal of Management in Engineering* , Volume 32, pp. 0-10.

Love, P., Davis, P., Ellis, J. & Cheung, S., 2010. Dispute causation: Identification of Pathogenic Influences. *Engineering, Construction and Architectural Management,* 17(4), pp. 404-423.

Love, P. & Edwards, D., 2004a. Determinants of rework in building construction projects. *Construction and Architectural Management,* 11(4), p. 259–274.

Love, P., Irani, Z. & Edwards, D., 2004. A Rework Reduction Model for Construction Projects. *IEEE Transactions on Engineering Management.*

Love, P. & Josephson, P., 2004. Role of Error-Recovery Process in Projects. *ASCE Journal of Management in Engineering,* 20(2).

Love, P. & Li, H., 2000. Quantifying the causes and costs of rework in construction. *Construction Management and Economics,* 18(4), p. 479–490.

Love, P., Lopex, R., Goh, Y. & Tam, C., 2011. What Goes Up, Shouldn't Come Down: Learning from Construction and Engineering Failures. *Procedia Engineering,* Volume 14, pp. 844-850.

Love, P., Mandal, P. & J Smith, J. G., 2000. *A Design and Construction Rework Minimisation Model Proceedings of the International Conference on Systems Thinking in Management,* Victoria, Geelong, Victoria: Deakin University.

Love, P., Mandal, P. & Li, H., 1999a. Determining the causal structure of rework in construction projects. *Construction Management and Economics,* 17(4), p. 505–517.

Love, P., Mandal, P., Smith, J. & Li, H., 2000. Modeling the dynamics of design error induced rework in construction projects. *Construction Management and Economics,* 18(5), pp. 567-574.

Love, P., Mohamed, S. & A Wyatt, A., 1997. Understanding rework in construction. Construction Process Re-Engineering. *Gold Coast,* Australia: Griffith University, pp. 269-278.

Love, P. & Sing, C., 2011. Determining the Probabilistic Distribution of Rework Costs in Construction and Engineering Projects. *Journal of Structure and Infrastructure Engineering,* pp. 1-13.

Love, P., Skitmore, R. & Earl, G., 1998b. Selecting a suitable procurement method for a building project. *Construction Management and Economics,* 16(2), p. 221–233.

Love, P. & Smith, J., 2003. Benchmarking, bench-action and bench-learning: rework mitigation in projects. *ASCE Journal of Management in Engineering,* 19(4), pp. 147-159.

Love, P. & Sohal, A., 2013. Capturing Rework Costs in Projects. *Managerial Auditing Journal* ,18(4), pp. 329-339.

Low, S. & Yeo, H., 1998. A construction quality costs quantifying system for the building industry. *International Journal of Quality & reliability Management,* pp. 329 - 349.

Maghrebi, M., Sammut, C. & Waller, T., 2014. Predicting the duration of concrete operations via artificial neural network and by focusing on supply chain parameters. *Building Research Journal* , 61(1), pp. 1-14.

Marosszeky, M., 2006. *Performance Measurement and Visual Feedback for Process Improvement. A Special Invited Lecture presented in the SMILE-SMC 3rd Dissemination Workshop,* Hong Kong: Centre for Infrastructure and Construction Industry Development of the University of Hong Kong.

Mastenbroek, Y., 2010. *Reducing rework costs in construction projects,* Published Bachelor's degree Thesis. Netherlands, University of Twente.

Mathi, K., 2004. *Key Success Factors for Knowledge Management*, Unpublised MBA Thesis. Germany, Internationales Hochschulinstitut Lindau, University Of Applied Sciences/ Fh Kempten, Germany

Meshksar, S., 2012. *Cost and Time Impacts of Reworks in Building a Reinforced Concrete Structure*, Unpublished MSc Thesis. Gazimağusa, North Cyprus, Eastern Mediterranean University.

Mills, A., Love, P. & Williams, P., 2009. Defect costs in residential construction. *ASCE Journal of Construction Engineering and Management*, 135(1), pp. 12-16.

Nesan, J. & Holt, G., 1999. *Empowerment in Construction: The Way Forward for Performance Improvement*, Hertfordshire. Somerset Research Studies Press.

Nylen, K., 1996. *Cost of Failure in a Major Civil Engineering* , Stockholm, Sweden, Division of Construction Management and Economics, Department of Real Estate and Construction Management, Royal Institute of Technology.

Oladinrin, T., Ogunsemi, D. & Aje, I., 2012. Role of Construction Sector in Economic Growth: Empirical Evidence from Nigeria. *FUTY Journal of the Environment*, pp. 50-60.

Orange, 2020. *Interactive Data Visualization.* [Online]
Available at: https://orange.biolab.si/home/interactive_data_visualization/
[Accessed 20 October 2020].

Oyewobi, O., Ibironke, O. T., Ganiyu, B. O. & Ola-Awo, A. W., 2011. Evaluating rework cost- A study of selected building projects in Niger State, Nigeria. *Journal of Geography and Regional Planning* , 4(3), pp. 147-151.

Oyewobi, L. & Ogunsemi, D., 2010. Factors influencing rework occurrence in construction: A study of selected building projects in Nigeria. *Journal of Building Performance*, The Institution of Surveyors Malaysia,1(1), pp. 1-20.

Palaneeswaran, E., 2006. Reducing rework to enhance project performance levels, Proceedings of the one day seminar on recent developments in project management in Hong Kong. *Centre for Infrastructure and Construction Industry Development*, Hong Kong, 5(10), pp. 1-5.

Patterson, L. & Ledbetter, W., 1989. The cost of quality: a management tool, Excellence in the Construction Project. *Proceedings of Construction Congress*, p. 100–105.

PMI, 2010. *A Guide to Project Management Body of Knowledge.* 4th ed. Project Management Institute.

Podges, T., 2017. *A decision model for the investment in technology to reduce concrete rework,* Unpublished MEng Thesis. South Africa, Stellenbosch University.

PRINCE2, 2017. *Managing successful projects with PRINCE2.* London: London (London) : TSO.

Prodonovich, S., Perry, L. & Taggart, A., 2014. Developing conceptual understandings of the capacity to aspire for higher education. *Issues in Educational Research,* 24(2), pp. 74-189.

Rhodes, B. & Smallwood., J., 2003. Defects and rework in South African construction projects. *Proceedings of the first CIDB Postgraduate Conference,* Port Elizabeth, South Africa, pp. 228-236. .

Rodrigues, A. & Bowers, J., 1996. System dynamics in project management: A comparative analysis with traditional methods. *System Dynamics Review,* 12(2), pp. 121-139.

Rogge, D., Cogliser, C., Alaman, H. & McCormack, S., 2001. *An Investigation into Field Rework in Industrial Construction,* Austin, Texas: No.RR153-11, Construction Industry Institute.

Rosowsky, D. & Stewart, M., 1996. *Quality Control and Factors of Safety for Reinforced Concrete Design,* s.l, Congress Report.

Solomons, W. J. (2015). *Investigating the impact of site activities and conditions on concrete quality of in-situ and precast construction methods,* Unpublished MEng Thesis. South Africa , University of Stellenbosch.

Sommerville, J. & McCosh, J., 2006. Defects in new houses; An analysis of data on 1 696 new UK houses. *Structural Survey,* 24(10), pp. 6-21.

Wasfy, M., 2010. *Severity and impact of rework, A case study of a residential commercial tower project in the eastern province-KSA,* Dhahran, Saudi Arabia, King Fahd University of Petroleum & Minerals.

Wowczko, W., 2015. A Case Study of Evaluating Job Readiness with Data Mining Tools. *International Journal for Infonomics (IJI),* 8(3).

Yates, J. & Lockley, E., 2002. Documenting and Analysing Construction Failures. *ASCE Journal of Construction Engineering and Management,* Volume 128(1). pp. 8-17.

Zhang, D., 2009. *Analysis of a Construction SmallProjects Rework Reduction Program for a Capital Facility*. Published MSc Dissertation, University of Waterloo, Ontario, Canada.

APPENDICES

A: Steel Reinforcement inspection sheet

EPC Contractor				Reference Number	
Project Name				Table Number	
Location				Page 1 of 1	
Working Items	Reinforcement bar installation				
Description		Unit	Standard		Actual data
Rebar Installation	Surface quality of rebar	/			
	Level of horizontal bar	mm	Deviation tolerance of ±2mm		
	Joint position	/	Should be staggered along the same horizontal plane		
	Rebar spacing	mm	Maximum tolerance of ±5mm		
	Vertical reinforcement bars	/	Rebars should be vertically upright		
Remarks (If any) Can be explained by attached drawing					
CONTRACTOR			EMPLOYER		
Civil Engineer			Civil Engineer		
Date			Date		

B: Formwork inspection sheet

EPC Contractor				Reference Number	
Project Name				Table Number	
Location				Page 1 of 1	
Working Items	Formwork				
Description		Unit	Standard		Actual data
Formwork installation and removal	Formwork material/type	/	Steel/wood		
	Embedded parts reserved holes	/			
	Elevation deviation	mm	± 5		
	Formwork clearance	mm	≤ 5		
	Formwork size		As per drawing		
	Axis displacement	mm	≤ 5		
Remarks (If any) Can be explained by attached drawing					

CONTRACTOR		EMPLOYER	
Civil Engineer		Civil Engineer	
Date		Date	

C: Slump test inspection sheet

<table>
<tr><td colspan="7">Concrete slump and temperature record sheet</td></tr>
<tr><td colspan="2">Date</td><td colspan="2"></td><td colspan="2">Record Number</td><td></td></tr>
<tr><td colspan="2">Unit Project</td><td colspan="2"></td><td colspan="2">Location</td><td></td></tr>
<tr><td colspan="2">Slump range</td><td colspan="2"></td><td colspan="2">Design strength</td><td></td></tr>
<tr><td colspan="2">Temperature range</td><td colspan="2">5 – 35 °C</td><td colspan="2">Concrete quantity</td><td></td></tr>
<tr><td>Item</td><td>Truck Number</td><td>Time</td><td>Temp (°C)</td><td>Slump (mm)</td><td>Blocks (pieces)</td><td>Remarks</td></tr>
<tr><td>1</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>2</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>3</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>4</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>5</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>6</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>7</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>8</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>9</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>10</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>11</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>12</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>13</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>14</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>15</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>16</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

D: Cast concrete inspection sheet

EPC Contractor			Reference Number	
Project Name			Table Number	
Location			Page 1 of 1	
Working Items	Concrete			

Description		Unit	Standard	Actual data
Cast concrete	Concrete strength	/	Steel/wood	
	Construction joint retention and treatment	/	Should comply with design requirements	
	Curing	/	Should comply with design requirements	
	Appearance quality	/		
	Axis displacement	mm	≤ 15	
	Verticality	mm	≤ 30	
	Elevation deviation	mm	± 30	

Remarks (If any) Can be explained by attached drawing

CONTRACTOR		EMPLOYER	
Civil Engineer		Civil Engineer	
Date		Date	

E: Non-conformance report

<table>
<tr><td colspan="2" rowspan="2"></td><td>Document Number</td><td></td></tr>
<tr><td>CAR Number</td><td></td></tr>
<tr><td>Section</td><td>Civil and structural</td><td>Effective date</td><td></td></tr>
<tr><td>Subject</td><td>Corrective Action Request form</td><td>Page Number</td><td></td></tr>
<tr><td colspan="4">Originator :</td></tr>
<tr><td colspan="4">1. Details of non-conformance/observations (finding statement)

Signed by Originator   Signed by Dept Rep</td></tr>
<tr><td colspan="4">2. Correction</td></tr>
<tr><td colspan="4">3, Investigation of root cause

Signed by Action Owner   Signed by Originator</td></tr>
<tr><td colspan="4">4. Corrective Actions

Signed by Action Owner ………………………………..

Signed by Originator ………………………………..</td></tr>
</table>

Target completion date …………………………………
5. Verification of effectiveness of corrective action   Signed by Action Owner ……………......... Date ………………………... Signed by Originator ……………......... Date ………………………..

F: Hazard Identification and Risk Assessment Template

Activity Number	Task	Identify Hazard	Identify Risk	Legal Requirements	Catergory	Likelihood	Severity	Inherent Risk	Controls	Applicable Safe Work Procedure (SWP)	Trainings Required	Monitering and Mearsurement Required	Likelihood	Severity	Inherent Risk	Control Owner

G: Questionnaire Survey

My name is Fionah Mazvita Mukondwa studying for an MEng in Construction Engineering and Management with the University of Stellenbosch. I am kindly inviting you to take part in my academic research which is investigating the extent to which various risk triggers contribute to rework during the construction of reinforced concrete structural elements. The aim of the study is to apply supervised machine learning in the formulation of a predictive data analytics model to minimize and eliminate rework during concrete construction.

1 Participant Information

1.1 Professional Background

- Contractor
- Consultant
- Site Engineer
- Site Foreperson
- Site Operative
- SHERQ Practitioner
- Other

1.2 Highest Level of Education

- Diploma
- Bachelors Degree
- Honours Degree
- Masters Degree
- Doctorate

1.3 Years of industrial experience

- 0 – 10 years
- 11- 20 years
- 21- 30 years
- >31 years

Based on your experience, answer the following questions regarding the extent to which the mentioned risk triggers contribute to rework prior to, during and after the placement of reinforced concrete.

2.0 PRE-PLACEMENT (Reinforcement works)

2.1 Insufficient rebar tensile strength

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

2.1 Incorrect rebar size changes

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

2.2 Poor surface quality rebars

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

2.3 Incorrect rebar placement and spacing

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

2.4 Incorrect cover to reinforcement

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

2.5 Inadequate overlap distance

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

2.6 Missing reinforcement bars

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

2.7 Loose ties

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

2.8 Un-staggered splices and connections

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

3.0 PRE-PLACEMENT (Formwork)

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

3.1 Deformed or damaged formwork boards

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

3.2 Wrongly dimensioned formwork.

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

3.3 Inadequately sealed formwork joints (not leak proof or /water tight).

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.5 Poor formwork board surface quality.

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.6 Incorrect formwork elevation.

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.7 Insufficient formwork support and bracing

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.0 DURING PLACEMENT

4.1 Using inappropriate poker head for vibrating

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.2 Erratic concrete delivery to site (cold joint formation)

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.3 Inadequate poking/vibration

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.3 Inadequate poking/vibration

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4. 5 Poor concrete workability

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.6 Over workability of concrete

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.7 Grout loss

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.8 Formwork displacement during placement

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

4.9 Placement under adverse weather conditions

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

5.0 POST PLACEMENT

5.1 Poor and inadequate curing

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

5.2 Damage during stripping

	0	1	2	3	4	5	6	7	8	9	10	
Insignificant	○	○	○	○	○	○	○	○	○	○	○	Very significant

H:Confusion Matrix Test Results

H 1: Confusion Matrix test results for dataset with 500 instances.

Actual \ Predicted	Expect Rework	Proceed to next stage	Σ
Expect Rework	1	9	10
Proceed to next stage	1	489	490
Σ	2	498	500

H 2: Confusion Matrix test results for dataset with 1 000 instances.

Actual \ Predicted	Expect Rework	Proceed to next stage	Σ
Expect Rework	18	13	31
Proceed to next stage	1	968	969
Σ	19	981	1000

H 3: Confusion Matrix test results for dataset with 2 500 instances.

Actual \ Predicted	Expect Rework	Proceed to next stage	Σ
Expect Rework	70	21	91
Proceed to next stage	14	2395	2409
Σ	84	2416	2500

H 4: Confusion Matrix test results for dataset with 5 000 instances.

		Predicted		
		Expect Rework	Proceed to next stage	Σ
Actual	Expect Rework	136	29	165
	Proceed to next stage	9	4826	4835
	Σ	145	4855	5000

H 5: Confusion Matrix test results for dataset with 10 000 instances.

		Predicted		
		Expect Rework	Proceed to next stage	Σ
Actual	Expect Rework	336	9	345
	Proceed to next stage	0	9655	9655
	Σ	336	9664	10000