Modeling the behavior of construction workers to predict their propensity for unsafe acts

By

Rafia Nawaz Awan

MS CE&M 00000321121

A synopsis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Construction Engineering and Management

Thesis Supervisor: Dr. Muhammad Usman Hassan

Department of Construction Engineering and Management

National Institute of Transportation (NIT)

School of Civil and Environmental Engineering (SCEE)

National University of Sciences and Technology (NUST),

Islamabad, Pakistan

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS thesis written by Miss. Rafia Nawaz Awan, Registration No. 00000321121, of CEM-11 (SCEE) has been vetted by the undersigned, found complete in all respects as per NUST Statutes / Regulations, is free of plagiarism, errors, and mistakes, and is accepted as partial fulfillment for the award of MS degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: ___________________________________

Name of Supervisor: Dr. Muhammad Usman Hassan

Date: ____

Signature (HOD): _____________________________

Date: ____

Signature (Dean/Principal): _____________________

Date: ____

Form TH-4

National University of Sciences and Technology

MASTER′S THESIS WORK

We hereby recommend that the dissertation prepared under our Supervision by:

(Rafia Nawaz Awan NUST 2019MSCEM00000321121)

Titled **Modeling the behavior of construction workers to predict their propensity for unsafe acts**

be accepted in partial fulfillment of the requirements for the award of **Master of Science** degree with (_____ Grade).

Examination Committee Members

1. Name: Dr. Khurram Iqbal Ahmad Khan Signature:_______________

2. Name: Dr.-Ing. Abdur Rehman Nasir Signature:_______________

3. Name: Dr. Saima Ehsan Signature:_______________

Supervisor's name: Dr. Muhammad Usman Hassan Signature:_______________

Date: __________________

____

Head of Department

Date: _________

COUNTERSIGNED

Date: ____________ Dean/ Principal

Certificate of Approval

This is to certify that the research work presented in this thesis, entitled " Modeling the behavior of construction workers to predict their propensity for unsafe acts" was conducted by Ms. Rafia Nawaz Awan under the supervision of Dr. Muhammad Usman Hassan. No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the Construction Engineering and Management Department in partial fulfillment of the requirements for the degree of Masters of Sciences in field of Construction engineering and Management.

Department of Construction Engineering & Management, School of Civil and Environmental Engineering. National University of Sciences and Technology.

Student Name: Rafia Nawaz Awan Signature: ________________

Examination Committee:

a) Dr. Khurram Iqbal Ahmad Khan Signature: ________________
(Assistant Professor, Dept. of Construction Engineering and Management, NIT, SCEE, NUST)

b) Dr.-Ing. Abdur Rehman Nasir Signature: ________________
(Assistant Professor, Dept. of Construction Engineering and Management, NIT, SCEE, NUST)

b) Dr. Saima Ehsan Signature: ________________
(Assistant Professor, Dept. of Psychology, Foundation University, Islamabad)

Supervisor Name: Dr. Muhammad Usman Hassan Signature: __________

Name of Dean/ HOD: Dr.-Ing. Abdur Rehman Nasir Signature: __________

Author’s Declaration

I, Rafia Nawaz Awan, hereby state that my MS thesis titled “Modeling the behavior of construction workers to predict their propensity for unsafe acts”, is my own work and has not been submitted previously by me for taking any degree from this National University of Sciences & Technology or anywhere else in the country/ world. At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student:
Rafia Nawaz Awan

Date:

Plagiarism Undertaking

I solemnly declare that the research work presented in the thesis title "Modeling the behavior of construction workers to predict their propensity for unsafe acts" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I undertake that if I am found guilty of any formal plagiarism in the above-titled thesis even after being awarded of MS degree, the University reserves the rights to withdraw/revoke my MS degree and that HEC and the University has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student/Author Signature:

Name: Rafia Nawaz Awan

This thesis is dedicated to my precious father who is and will always be my living, breathing superhero.

ACKNOWLEDGMENTS

Countless gratitude to Almighty ALLAH, Who is the omnipotent and the omnipresent, for blessing me with the chance, health, courage, strength, patience, and knowledge to complete my research work.

I would like to express my sincere gratitude to my research supervisor Dr. Muhammad Usman Hassan, whose countless inspirations and guidance made it possible to complete this research. His patience, motivation, encouragement, immense knowledge, and meticulous editing helped me throughout the research. I am extremely grateful that he took me on as a student and continued to have faith in me over the years. I genuinely appreciate the valuable time and personal support accorded by him.

I am also thankful to the committee members, Dr. Khurram Iqbal Ahmad Khan, Dr. Saima Ehsan, and Dr. Abdur Rehman Nasir for their help and guidance.

In the end, I would like to thank my parents, siblings, and my friends for their absolute support and continuous encouragement.

ABSTRACT

The construction sector is leading in the number of accidents caused by the unsafe behavior of workers. Unsafe behavior can arise from a worker's personal preferences or from any external unsafe condition in his working environment. Personal preferences are the intentional unsafe practices of workers and various behavioral constructs drive these preferences. On the other hand, the absence of no measures and controls for unsafe behaviors at the organizational level can be treated as an external unsafe condition that leads to unsafe acts. Therefore, this study proposes a propensity prediction engine powered by the classification algorithm of Artificial neural networks (ANN). The ANN-based propensity model takes quantified values of individual features of behavior-modifying constructs as inputs and provides outputs in the form of classification of workers as safe or unsafe. The behavioral constructs are taken from the theory of planned behavior (TPB) and the individual features of these constructs are explored from previous studies and field surveys. The model is a multi-layer feed-forward network with back-propagation built on an architecture of 10-16-6-2 and has been trained, validated, and tested using Keras API of Tensorflow. The study also presents a framework for practical implementations of propensity prediction engine for construction organizations. Specialized behavior interventions are proposed to be included in safety training programs for worker's classified as unsafe by the prediction engine. The engine will help construction organizations in improving their safety training program by providing a way of managing behabehavioragement gaps at the organizational level.

KEYWORDS: Unsafe behavior, Unsafe acts, Theory of planned behavior, Artificial neural networks, Behavior prediction, Safety training programs

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS, ABBREVATIONS and acronyms

SCM : Swiss Cheese Model

HFACS: Human Factor Analysis and Classification Systems

PSM: Proactive Safety Management

TPB: Theory of Planned Behavior

TRA: Theory of Reasoned Action

SA: Safety Attitude

PN: Perceived Norms

PBC: Perceived Behavioral Control

SEM: Structural Equation Modeling

BBS: Behavior-Based Safety

ML: Machine Learning

DT: Dicision Trees

SVM: Support Vector Machines

RF: Random Forests

ANN: Artificial Neural Networks

CSS: Classification Criteria Sheet

API: Application Programming Interface

Chapter 1

Introduction

1.1. Background

The construction sector is leading in the number of accidents and fatalities. Accidents cause several direct costs such as the cost of injuries and can cause indirect damages such as psychological outlays for the workers (Tixier *et al.*, 2014). Construction worker's safety is seen as a complex phenomenon. Risky situations are always present on construction sites because of many outdoor activities, work at heights, complicated site layouts and equipment operation procedures (Choudhry and Fang, 2008). The construction industry is also dynamic in nature as many vigorous challenges regarding work hazards are encountered, which are further aggravated by organizational and personal characteristics. In a hazardous industry like construction, unsafe behavior among workers seems to be a critical factor causing workplace accidents (Fogarty and Shaw, 2010). Many investigations of construction accidents have shown that workers' unsafe behavior is one of the common causes of accidents. Around 80% of all construction site accidents are caused by unsafe behaviors of employees (Shin, Gwak and Lee, 2015). Studies on construction safety have also identified several personal, organizational, and environmental factors which influence behavior of workers (Huang *et al.*, 2016). According to Brown, Willis and Prussia (2000), the causes of unsafe acts of workers can be categorized into three basic themes: person as a cause, system as a cause and system–person interrelationships. The person as cause theme sees employee behaviors as the most critical precursors to unsafe acts and accidents.

In construction safety research, fragmented factors which lead to unsafe acts have been explored widely. These fragmented factors have been classified as human factors which include safety behavior, safety attitude, risk tolerance (Guo, Yiu and González, 2016; Wang, Zou and Li, 2016), and organizational/environmental factors which include senior management's commitment and frontline supervision (Zou and Sunindijo, 2013; Fang and Wu, 2015). Amongst prevailing classification methods, a particularly comprehensive approach is Human Factors Analysis and Classification System (HFACS). HFACS was initially devised within the aviation field for accident investigation and analysis (Wiegmann and Shappell, 2001). The HFACS is based on Swiss Cheese model (SCM) by Reason (1990). HFACS illustrates four stages of failure which are as follows: unsafe acts, preconditions for unsafe acts, unsafe supervision, and organizational influences. The unsafe acts (1st level) is closest to accident (Cohen, Wiegmann and Shappell, 2015). An unsafe act can have multiple preconditions which may include worker's psychological and physiological state or the presence of a an unsafe condition (organizational factor), paving way for an accident (Xia *et al.*, 2018). According to Chi, Han and Kim, (2013) accidents occur when an unsafe behavior meets an unsafe condition. In order to eliminate the chances of accidents, firstly, all unsafe conditions (organizational factor) have to be removed. The organizational factors e.g., senior management's commitment and effective supervision, can be improved if the organization selects its employees by scrutinizing them on their overall safety behavior (Chi, Han and Kim, 2013). This approach provides a solution for eliminating both the unsafe conditions and unsafe behavior. The organization not assessing its workers for their safety behavior is providing an unsafe condition in advance to its employees (Zhou, Fang and Mohamed, 2011). Once the organization decides to assess worker's behavior, it can strengthen its overall safety environment.

1.2. Research Problem

Construction is globally known as an industry with the greatest hazards. Despite the significant research effort done to eliminate the identified causes of accidents, the construction sector is still observed as amongst the most hazardous workplaces in the world. This is because of the adaption of reactive approach of safety management which addresses the problem after an accident has occurred. This reactive approach has failed in reducing accidents caused by unsafe acts of workers, since these unsafe acts are a dominant cause of workplace accidents. Although, in recent years the concepts of safety culture and unsafe environment is being tossed and practiced by many multi-national companies but the measures to identify and control unsafe acts are still defective. This is because the safety systems are introduced after an incident occurs and safety managers are unable to forsee the accidents associated with the individual unsafe acts of workers. Once an accident happen, the corrective measures are taken that include the work related safety trainings. Figure 1.1 shows that how the reactive approach identifies a problem after an incident occurs whilst the proactive approach identifies the problem, assesses it and devises methods to counter it. In reactive approach, the root cause of accident is ignored and the focus is diverted towards how the injured worker can be compensated through insurance.

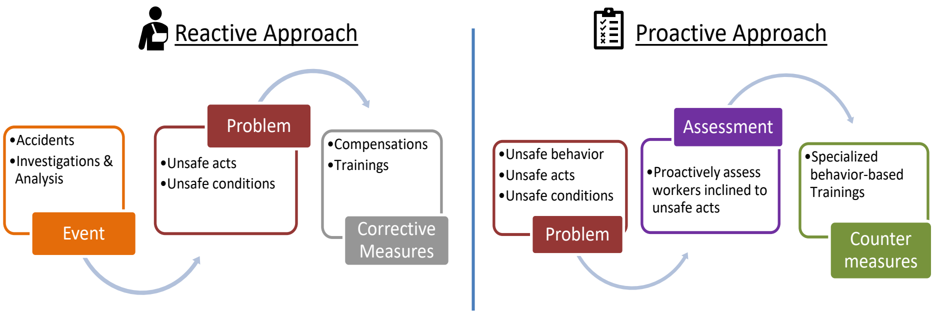

Figure 1.1: Comparison between reactive & proactive systems

Most of the times, the accidents caused by unsafe behaviors are not investigated enough in order to avoid the victim-blaming situation since the worker is already suffered an injury. This results in tolerance of unsafe behaviors at management level and can be identified as a failed defence of safety management against a mishap. Thus, the proactive safety systems are needed, that can assess and predict the behavior of workers so that the unsafe workers can be given more training or education regarding the importance of safety on job-site. One way to counter individual unsafe acts is to understand worker's behavior towards safety and the factors influencing it. Therefore, this study postulates that assessing worker's proneness to unsafe acts can help management in effectively managing individual unsafe acts as a proactive safety management system (PSM). The proactive modelling approach will help safety personals in assessing overall behavior of a worker proactively, so as to avoid unsafe acts caused by personal characteristics.

1.3. Previous studies

After successful application in Aviation industry, HFACS has been verified as a tool for human error investigations in some other fields, such as chemical industry, railways, mining, oil & gas and healthcare (Reinach and Viale, 2006; Cohen, Wiegmann and Shappell, 2015; Theophilus *et al.*, 2017; Rostamabadi *et al.*, 2019; Yıldırım, Başar and Uğurlu, 2019). Recently, the framework was also applied in the construction industry and modified HFACS models are proposed through the empirical research (Xia *et al.*, 2018; Tang *et al.*, 2021). Furthermore, previous studies, based on the causes of construction accidents have focused on unsafe site conditions e.g., defective tools and devices, unguarded openings, and improper storage of equipment and materials. The behavior of worker towards safety as a precursor of unsafe acts is yet to be defined in HFACS since in construction safety management, identifying and understanding unsafe behavior is seen as an important aspect when considering the overall safety culture of organization. The behavior prediction models have been developed and validated by several researchers (Seo, 2005; Cui *et al.*, 2013; Fang, Wu and Wu, 2015; Guo, Yiu and González, 2016) and have offered important understandings about safety behavior, its cognitive and affective mechanisms along with influencing factors. Additionally, from the area of psychology, Theory of planned behavior (TPB) (Ajzen and Madden, 1986) and Theory of reasoned action (TRA) (Ajzen and Madden, 1986; Fishbein and Ajzen, 2011) are adopted by various safety researchers for explaining and predicting behavior (Johnson and Hall, 2005; Fogarty and Shaw, 2010; Goh and Binte Sa'adon, 2015; Fang, Zhao and Zhang, 2016; Goh *et al.*, 2018; Xu, Zou and Luo, 2018). The research in the past have been using traditional modelling techniques from statistical sciences for instance linear and logistic regression or structural equation modelling (SEM) to test behavioral models (Fogarty and Shaw, 2010). Machine learning techniques are setting a new growing trend in construction safety research. It is subset of artificial intelligence and is defined as an algorithmic approach which instead of rule-based coding/programming, learns from data and improve its performance. Machine learning can prove to be a more accurate and useful alternative to traditional data modelling (Breiman, 2001). Among the machine learning techniques , Artificial Neural Networks are widely implemented in many construction safety studies (Goh and Chua, 2013; Goh and Binte Sa'adon, 2015; Patel and Jha, 2015b; Fang, Zhao and Zhang, 2016a; Goh *et al.*, 2018).

1.4. Gap in Previous Studies

Although, over the past few years, behavioral modelling has become one of the popular research topics in the domain of construction safety and has invoked growing research attention. Distinctly, Goh and Binte Sa'adon, (2015) did effective research in proposing an ANN model for construction workers unsafe behavior at height and reported valued findings. Similarly, Patel and Jha, (2015b) has used safety climate factors to predict unsafe behavior of workers. However, their study lacked detailed validity assessments and consideration of the aspect of behavioral factors which result in unsafe acts. So far, however, there is an absence of a well-constructed, reliable, effective, and comprehensive predictive model, for the assessment of behavior of construction workers. Even though research on unsafe behavior prediction is available for other sectors, its results cannot be applied to construction as they emphasis on specific sector-related kinds of determinants of safety. Also, construction is exclusively characterized by its dynamism where type of work, situations and their consequential risks are constantly shifting. Therefore, this area requires explicit research because using the findings from research of other sectors can be misleading. Now, further safety improvements require the construction industry to pay more devotion to eliminating unsafe acts.

This research aims on exploring personal factors which influence behavior of construction workers and proposes an assessment criterion for organizations to effectively assess their workers, hence, reducing the number of unsafe acts on construction sites,To address the above-mentioned conundrum of safety behaviors and unsafe acts, this study focuses on proposing and validating an Artificial neural network model for predicting unsafe behavior of workers at construction sites. The research adopts the theory of planned behavior to evaluate factors for behavior modelling. It uses the Swiss cheese model (SCM) to explain the problem of individual unsafe acts and proposes a modified HFACS with the addition of behavioral factors for human error analysis. The purpose of adding and modifying HFACS is to use it for proactively removing human-based errors which cause accidents. This multidisciplinary research combines tactics from the fields of engineering, psychology and sociology and contributes to the research in safety domain of construction sector. The proactive prediction of unsafe behavior has the potential to minimize workplace accidents and will consequently be very useful in safety improvement on construction sites.

1.5. Research Questions

For addressing the research problem, it is important to understand the research questions which can act as drivers in order to conduct the research in an effective way. For proposing a solution of a problem, it is evident to understand what actually the problem is, and what and where we can find the possible pieces of information which can help us in delivering the solution.

Following are the research questions which have directed this study.

i. How can we obtain worker's behavior prediction to make informed decisions on hiring, safety training and interventions?
   a. Can we predict a worker's propensity to unsafe acts using his distinctive behavioral features?
   b. What are the behavioral features that can be used to predict behavior?
   c. How can we model behavior by using machine learning algorithms?
   d. Can we design a tool to proactively cater behavior of workers by predicting their proneness to unsafe acts?

1.6. Research Objectives

The objectives of this research are:

i. To investigate individual features influencing the behavior of construction workers.
ii. To develop a propensity prediction model for identifying construction workers that are prone to unsafe acts.
iii. To propose a framework for dedicated training programs integrated with modified incident analysis model for continues improvement in behaviour of workers

1.7. Organization of thesis

The thesis contains five chapters. An outline of each chapter is provided below.

1. ***Chapter 1- Introduction***: It includes brief introduction of study followed by problem statement, research questions, research gap, and research objectives.
2. ***Chapter 2- Literature Review:*** It covers an overview of behavioral studies in construction safety research. It provides a comprehensive review of accident causation models, theories of behavior, factors influencing behavior and use of artificial neural networks (ANN) in construction safety research.
3. ***Chapter 3- Research Methodology***: This chapter includes a complete research methodology of the study. The research methodology consists of three major stages.
    i. Stage 1 is the literature review which is further divided into two steps i.e., research gap identification, research problem and formulating the research objectives. Step 2 includes detailed literature review of

behavior predicting theories, human error models, influencing factors, and modelling techniques.

ii. Stage 2 includes the formation of an ANN model for prediction of unsafe behaviors leading to unsafe acts.

iii. Stage 3 entails validation of proposed ANN model through a propensity prediction engine.

4. ***Chapter 4- Results and Analysis:*** This chapter has the detailed analysis of ANN predictive modelling approach used in the research. It includes the literature review and field surveys conducted to get the inputs of model. The ANN Architecture, Analysis and predictive accuracy of ANN model and modification of HFACS are discussed.

5. ***Chapter 5 – Conclusions and Recommendations:*** The final chapter contains the conclusions of research work; limitations of the study are discussed, and recommendations are proposed for future studies.

Chapter 2

Literature review

Overview

The literature review includes all relevant literature for a thorough understanding of all concepts and techniques used in this study. These topics include literature about causes of human errors leading to unsafe acts, theories of predicting and modifying behavior, variables which influence the main antecedents of behavior and behavior modelling techniques. Figure 2.1 shows the an overview of the topics explored in this chapter.

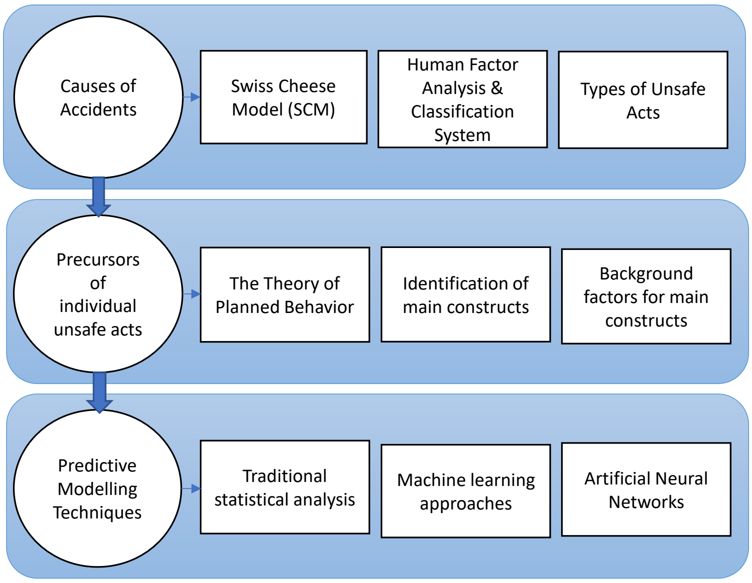

Figure 2.1: Schematic diagram of literature topics

2.1. Safety management in construction industry

Construction, being one of most hazardous industry, is suffering from many accidents which result in deaths, injuries and illness of workers along with various direct and indirect losses (Choudhry, Fang and Mohamed, 2007). Internationally, construction industry holds much important as it provides huge economic outputs by underpinning a variety of economic events and helps in achieving the social and fiscal objectives of countries (Lingard, Cooke and Blismas, 2012). Regardless of its significance, construction industry is among topmost unsafe industries. The accident and fatality rates of construction industry are significantly higher than the average of all other industries. Thus, it is essential for construction industry to control and improve its safety management systems (Zaid Alkilani and Jupp, 2013). The construction industry stakeholders usually focus on schedules, cost and quality of projects and safety is discussed at last. Safety concerns are considered less important and are put in the back seat during all phases of construction projects. Employers focus on maximizing their projects in limited time and ignore the need for the establishment of thorough and effective policies for accident prevention. The inability to foresee the actual high costs of an accident results in less attention given to the safety (Hollnagel, 2002; Mitropoulos, Abdelhamid and Howell, 2005). Therefore, it is important to understand the high stakes associated with high costs of injuries and effective safety management systems should be devised to proactively foresee and prevent accidents.

Different approaches have been proposed in various research articles for improving construction jobsite safety. The most common approaches implemented in several safety programs are as follows:

2.1.1. Behavior-based safety (BBS)

It improves safety performance of employees by keeping a checklist of safe or unsafe practices and observing and inspecting employees and rewarding or punishing workers on the basis of their behaviors (Choudhry, 2014; Guo, Yiu and González, 2016; Xia *et al.*, 2017; Curcuruto and Griffin, 2018).

2.1.2. People-based safety (PBS)

It improves workplace safety by centering more on people-based factors which influence employees' attitude and safety culture of organizations (Wiegand, 2007).

2.1.3. Cultural intervention

This approach improves workplace safety by reducing workers' unsafe behaviors through interventions based on safety awareness, climate and attitude (Mohamed, 2002; Choudhry, Fang and Mohamed, 2007; Mohamed, Ali and Tam, 2009).

This study has adopted the PBS approach, as it focuses on personal factors which act as precursors of an unsafe act and influence worker's overall behavior.

2.2. Causes of Accidents in Construction Industry

Accidents are caused by various factors. It is not right to say that they just happen. Both unsafe conditions and unsafe acts along with many other facilitating factors, can cause an accident (Baldissone et al., 2019). Accident causation models are methods or frameworks which help in identifying root causes of accidents. (Hamid, Majid and Singh, 2008; Wu, Gibb and Li, 2010). There are various accident causation theories i.e., Domino Theory put forward by Heinrich in 1930, Multiple Causation Theory proposed by Petersen in 1971 and Human Error theory presented by Abdelhamid in 2000 (Mitropoulos, Abdelhamid and Howell, 2005; Poor Sabet, 2013). Even if unsafe conditions are not present, humans often intentionally decide to go for an unsafe act, which can be classified as violations from rules or individual unsafe acts (Fogarty and Shaw, 2010). Generally, the main purpose of human error theory is to look for better designed workplaces and tools which cater human limitation. Other human error models such as human factor model, Ferrel theory and various behavior models, have also been derived from this theory. (Mitropoulos, Abdelhamid and Howell, 2005).

2.2.1. The Swiss Cheese Model (SCM) of Accidents Causation

This model was originated in 1988 when a psychologist James Reasons was writing a book named as Human Error (Reason, 1990). This model says that the accidents do not happen due to a single reason. They happen when a system's defense wall against an accident has a fault in it and it fails to provide protection against a problem. These faults are represented as holes in a cheese model. Figure 2.2 represents a SCM representing how loopholes at different levels of a system or organization lead to an accident. These unchecked, unidentified holes from each defence level of organization result in an accident. The model explains that unsafe acts are closest to mishaps and the hazard travels through each defence wall in a linear manner resulting in an accident at worker's level. In order to avoid a mishap, the model suggests checking each level of a system and identify possible loopholes and cater them effectively to break the chain of events. According to this model, if we close the loopholes at the first and highest level, i.e., organizational level, the chain of loopholes is closed once for all levels underneath (Reason, 1990).

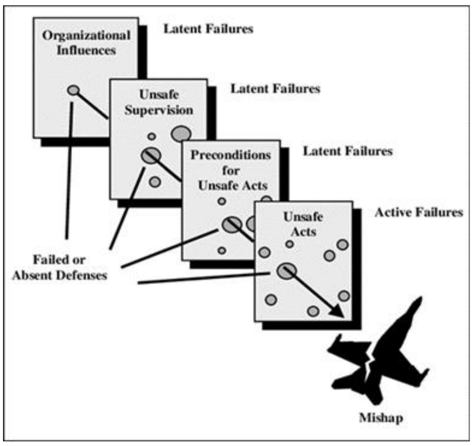

Figure 2.2: Reason's (1990) Swiss Cheese Model

According to Larouzee and Le Coze, (2020), this model has been adopted in almost every industrial field and is the most cited accident causation model. It has been used in industries like aviation, marine, healthcare, nuclear, defense, oil and gas, traffic, rails and road. This study is concerned with closing the loophole of regarding the tolerance of unsafe behaviors at organizational level, so that the hazard can be stopped from causing an incident at active individual levels.

2.2.2. The Human Factors Analysis and Classification System

HFACS was first proposed by Shappell and Wiegmann, (2000) for analysis of aviation accidents. The model is based on James Reason's SCM. According to Shappell and Wiegmann, (2000) if accidents are to be reduced, it is crucial to focus more on human factors, as they are the most common cause of accidents (Wiegmann et al., 2004; ElBardissi et al., 2007). In the swiss cheese model, Reason (1990) suggested that incidents take place when several pieces of a system do not interact with each other effectively. Basically, the failure to interact is seen as "holes" in different sheets of "cheese", and these holes then lead to an accident. There are four levels in Reason' s model which describe latent and active failures. These levels are: a) unsafe acts, b) preconditions of unsafe acts, c) unsafe supervision, d) organizational influences. The HFACS is designed with same four levels as proposed by Reason (1990). But they included 19 categories as causes of accidents within the four main levels of model. Figure 2.3 displays the actual model proposed by Wiegmann and Shappell (2001). Safety violations under unsafe acts, in the fourth level of HFACS can be called "breaking the rules," Violations can be habitual and are often strengthened by poor safety management system i.e., management tolerates safety violations which results in a poor safety compliance in frontline workers (Cohen, Wiegmann and Shappell, 2015). These violations are the individual unsafe acts of workers caused by the personal preferences of workers.

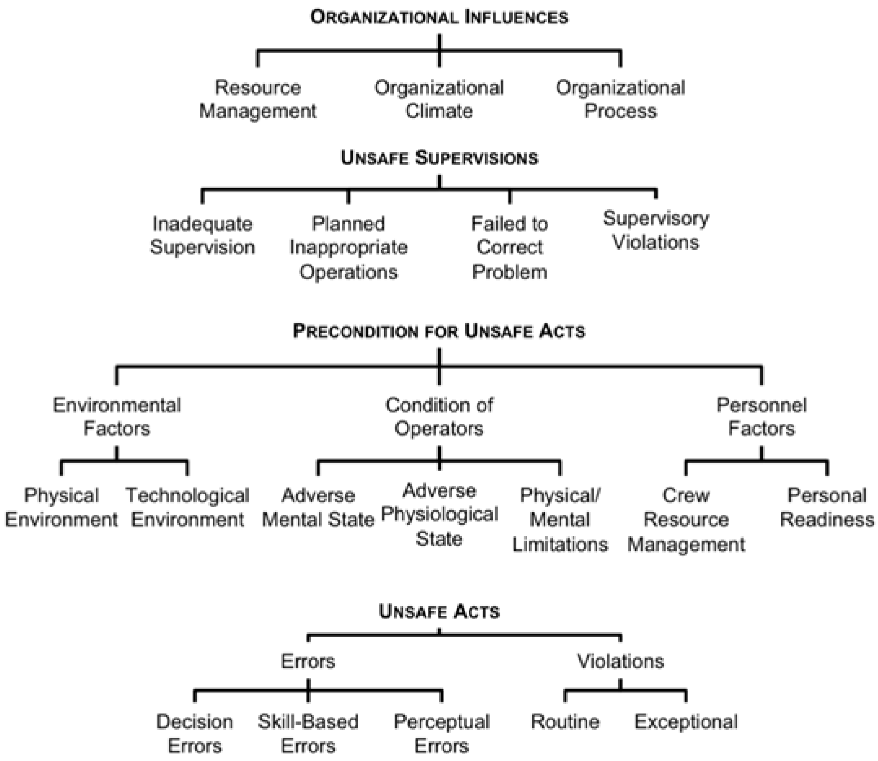

Figure 2.3: HFACS by Shappell & Weigmann (2001)

Therefore, this study proposes a model for proactively predicting and assessing whether a worker will opt for safety violation or not. This will help in reducing unsafe acts on sites and will try to close the "holes" at organizational influence level, henceforth, closing the main gate which leads to accidents. Various studies incorporating HFACS framework are listed down in Table 2-1.

Table 2-1: HFACS based research work

Field/Industry	**Research articles**
Aviation	(Shappell and Wiegmann, 2000; Wiegmann *et al.*, 2004; ElBardissi *et al.*, 2007; Cohen, Wiegmann and Shappell, 2015)
Oil & Gas	(Theophilus *et al.*, 2017; Uğurlu *et al.*, 2018; Rostamabadi *et*

	al., 2019)
Mining	(Yıldırım, Başar and Uğurlu, 2019)
Railways	(Reinach and Viale, 2006; Baysari, McIntosh and Wilson, 2008)
Construction	(Garrett and Teizer, 2009; Xia *et al.*, 2018; Ye *et al.*, 2018; Tang *et al.*, 2021)

2.2.3. HFACS for Construction Industry

In general, HFACS is rarely applied in construction projects for accident analysis (Garrett and Teizer, 2009). For HFACS to effectively work in construction industry, its original form and classifications levels need adjustments according to the features of construction field.

The changes in HFACS framework can be made by adding new domain specific item or layer or by deleting redundant items or layers (Uğurlu et al., 2018). Recently, Xia *et al.*, (2018) have modified the original HFACS framework for construction and added a new level named as "Environmental Influence". Figure 2.4 shows the modified HFACS for construction industry. The new level describes the environment influences which are external to the project. The items included in the level are identified as insufficient coordination between project stakeholders, weak social environment and improper or lack of regulations and enforcement (Xia *et al.*, 2018).

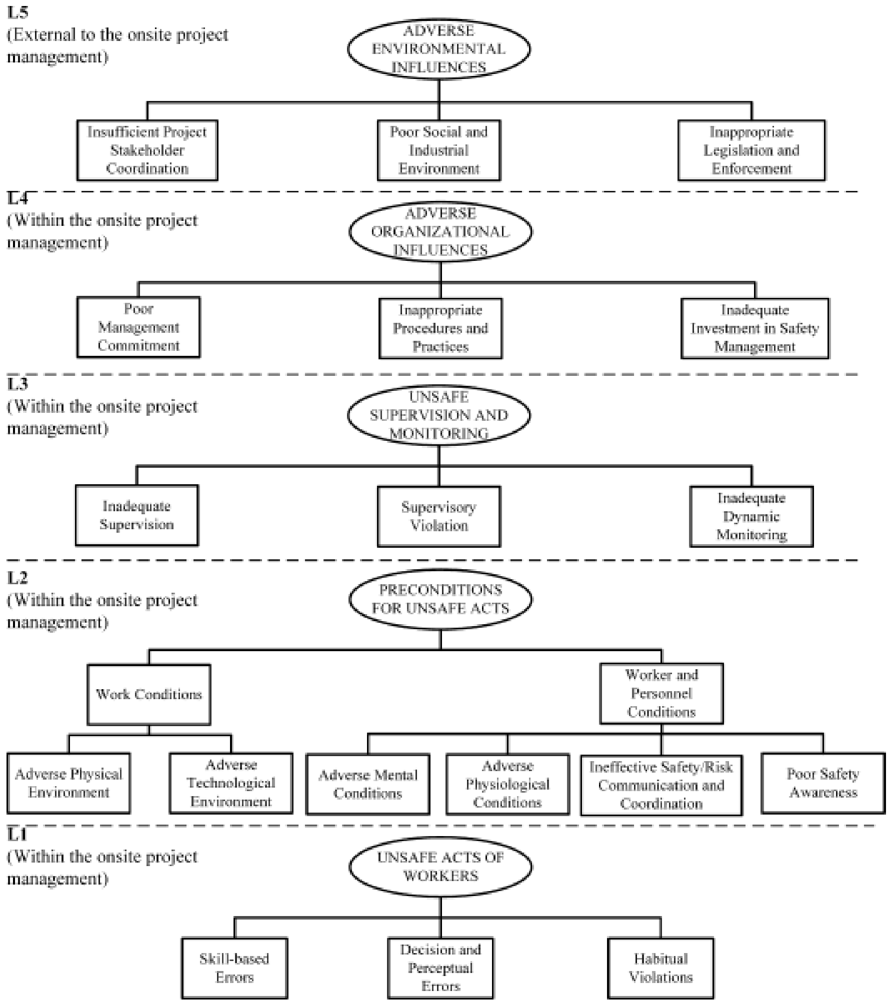

Figure 2.4: HFACS for construction industry

2.2.4. Types of unsafe acts

Violations in the new construction HFACS framework are described as workers' intentional deviations from safety rules and regulations, for example, not wearing PPE, chatting, eating, or smoking during work hours, and ignoring safety protocols of technical processes. Usually, in the other hazardous jobs, the rate of safety violations is less than rate of errors, but severity of violations is primarily high (Fogarty and Shaw, 2010). But in the construction industry, both the frequency and severity of violations in high. Habitual violations mainly occur over a very long period of either neglecting or not knowing the regulations or appropriate skills. The workers do habitual violations in almost every project they work on, most of the time the supervisors or senior management had tolerated the violations. This act of ignoring a violation makes a work environment prone to accidents (Fogarty and Shaw, 2010; Ye *et al.*, 2018). This implies that it is essential to create such conditions can reduce the likelihood of habitual safety violations. On the other hand, exceptional violations are also encountered on sites which are accidental violations. Such violations are difficult to predict and seldom occur. They occur according to the work processes, working scenarios and many other dynamical features. Sometimes, the scholars also call such violations as “best practice” of workers according to the situation (Aniekwu, 2007). Also, they are unexpected, so it is impossible to deal with them proactively.

Therefore, this study omitted the exceptional violations and focuses on habitual violations only. Safety violations are defined as an unsafe act which is a deliberate deviation from pre-defined rules of save work processes while errors are unintended consequences caused by mistakes made by individuals (Reason et al., 1990). Henceforth, the term safety violations in the research refers to habitual violations by workers and are called as individual unsafe acts of workers. Also, the adverse environmental influences have also been kept constant since we are looking from the perspective of strengthening organizational safety environment. Therefore, it is postulated that organization’s safety management should be strong enough to not provide a “hole” in their safety system which can lead to an unsafe act by worker at level 1. The preconditions of individual unsafe acts are further explored through the literature review.

2.3. Precursors of individual unsafe acts

Previous research has reported various organizational factors and socio-cultural factors that indirectly introduce unsafe acts in workers. There are four main factors which are considered to be the precursors of unsafe acts. These main factors are individual factors, job factors, management factors, and workgroup factors (Aksorn and Hadikusumo, 2007). These factors include various sub-factors. The individual factors can further be classified into psychological and physiological factors. These include lethargy, lack of motivation, attitude, uncomfortable while doing work, stress, drug abuse, doing work in hurry, macho syndrome, and overconfidence (Clarke, 2012). The job factors include site conditions, time pressure, difficult task and productivity pressure (Choudhry, 2014). Similarly, management factors factors include supervision style, pressure, safety protocols, the fear of penalty, rewards etc,. Workgroup factors include sub-factors such as coworker's influence, teamwork and influence of the overall attitude of worker's workgroup. (Aksorn and Hadikusumo, 2007; Khosravi *et al.*, 2014). All these factors contribute towards unsafe behavior of workers, leading to unsafe acts. It is also argued unsafe conditions also facilitate an unsafe act. For instance, a bad weather condition combined with worker's carelessness can lead to an accident. Therefore, while studying unsafe acts, it is important to look for unsafe conditions, which can be the drivers of unsafe acts of workers (Baldissone *et al.*, 2019). All these above mentioned factors combine up to create worker's behavior towards unsafe acts.

2.3.1. Worker Behavior

Behavior is the way in which a person acts. It is the action against a particular response or a situation (Liu et al. 2019). At construction sites, workers encounter various challenging situations, like conflicts, accidents, and disagreements. Different workers react differently in such situations, which not only impacts their productivity and safety but hampers others' performance as well. Therefore, many researchers have studied workers' behavior at construction sites through supervisor's intervention (Fang, and Wu, 2015), use of behavioral experiments and surveys (Choi et al. 2017), and hazard recognition (Namian et al. 2016). All such researchers have examined the behavior of the workers with the help of experts' opinions, or through experimental setups, and compared the results with the performance of the workers.

2.3.2. Safe vs unsafe worker characteristics

Characteristics of a worker are determined by the outcomes of his behavior. These results of these outcomes is either a safe or unsafe act and can be classified as safe or unsafe. The unsafe outcomes include accidents and incidents that are incurred either by a worker or other related workers (Hamid, Majid and Singh, 2008). There are many factors which contribute towards safe and unsafe acts of workers. A study by Wang, Zou and Li, (2016) has reported the relationship between risk tolerance of workers and stated that four factors which include subjective perception, work experience, work knowledge, work characteristics and safety management system will define a worker's behavior on construction sites. These factors can be called as characteristics and are mostly related to mental and cognitive process of workers. The way in which a worker's mental process comprehends all these factors will distinguish a worker's behavior from other. Whether the behaviour is safe or unsafe depends upon the outcomes of the behavior (Shin *et al.*, 2014). Other characteristics include personality traits which lead to habitual unsafe behavior. Researchers have also studied the relationship between various personality traits which can be put forward as characteristics of workers (Ma, Guo and Fang, 2021). Similarly, another study takes into account the influence of risk taking behavior on actions of workers with respect to age of workers (Man, Chan and Wong, 2017). Classifying worker's as safe and unsafe based on their mental process is a difficult task and may require a cutting edge technology as reading each human mind and predicting their actions can be an impossible task (Fang, Zhao and Zhang, 2016a).

An alternate approach adopted by Goh et al., (2018), of classifying whether a worker is safe or unsafe, takes the frequency of unsafe acts performed by workers. This approach comes under behavior-based safety, and classifies safe and unsafe workers using a rating scale. A predefined list of unsafe acts is prepared and workers are observed for a specific time period. Then the number of times a worker does an unsafe act, as listed in predefined list, is counted and converted into a percentage. If the percentage of unsafe acts is below a certain threshold, the worker is categorized as safe, otherwise unsafe (Goh *et al.*, 2018).

2.3.3. Propensity of unsafe act

The propensity of unsafe act can be defined as an individual's tendency of making an erroneous decision of doing an unsafe acts (Huang *et al.*, 2016). The worker's deliberate decision of doing an unsafe act, which he may or may not know, can lead to an accident. Most of the times, the workers donot regard minor deviations from safety protocols as unsafe acts and take the minimal consequences of such acts as of no significance. The unsafe behavior is one of the reasons for this individual tendency of unsafe acts. Various individual level factors such as risk perception, hazard perception, safety knowledge , safety motivations and accident experience etc., contribute towards unsafe behaviors and effect a worker's propensity of unsafe acts. (Shin, Gwak and Lee, 2015; Huang *et al.*, 2016). The study aims to further explore the individual level factors for making a behavior prediction model that can predict the propensity for unsafe acts.

2.4. Behavior prediction

Behavior prediction is usually done with either personality tests, through a set of predefined characteristics or otherwise it is evaluated by a certified psycologist. There are also various tests and theories of behavior prediction, widely used in research works of many industries. A few are discussed below.

2.4.1. Test for behavior prediction

Some studies have predicted unsafe behavior of the construction workers by analuzing their personalities. The Myers-Briggs Type Indicator (MBTI), Eysenck Personality Questionnaire, Cattell 16 Personality Factor (16PF) test, and the Big 5 are frequently used personality tests. These tests are based on trait theory and they can describe a simple dimension of an individual's personality. Hasanzadeh et al. (2018) used the Big 5 personality test to determine that workers with high neuroticism were more easily distracted and lost attention in work. Sing et al., (2014) used the Eysenck Personality Questionnaire (EPQ) to determine that bar benders with high psychoticism had a higher risk of accidents. A very recent study of Ma, Guo and Fang (2021), has used the Myers-Briggs test to predict unsafe behavior of construction worker's on bridge construction project (Ma, Guo and Fang, 2021).

2.4.2. Theories of behavior prediction:

There are two most prominent theories being adopted by researchers for human behavior predictions. These are "The Theory of Reasone Action (TRA)" and "The Theory of Planned Behavior".

TRA**:** In the 1970s, Fishbein and Ajzen developed the theory of reasoned action (TRA), with an attempt toidentify a set of variables that can account for a substantial proportion of the variance in any given behavior (Fishbein 2008) (p.834) (Ajzen and Madden, 1986). The old version of TRA proposed that the behavior is a determined by the attitude along with the personal and subjective norms with intentions acting as mediator. The perceived behavioral control is added as an additional factor after extensive studies and experiments.

TPB: The theory of planned behavior is a modified form of theory of reasoned action (Ajzen and Madden, 1986; Fishbein and Ajzen, 2011). It introduces a small number of variables that have the ability to explain the variance in a particular behavior. The theory suggests that an individuals' behavior can be predicted from his/her intentions and intentions are formed through three underlying constructs. These constructs include attitude towards behavior, the perceived norm, and the perceived behavioral control (Ajzen and Madden, 1986).

2.4.3. Constructs from theories

According to TPB, a belief is a personal prospect about an object that has certain attributes (Fishbein and Ajzen, 2011). Attitude is defined as "*a latent disposition or tendency to respond with some degree of favorableness or favorableness to an object, person or event*" (Ajzen and Madden, 1986). Perceived norm is defined as perceived social pressure to perform a particular behavior. PN consists of descriptive norms which depict the wishes and the actions of important referent persons. PBC is defined as "*the extent to which people believe that they are capable of performing a given behavior, that they have control over its performance*"(Fishbein and Ajzen, 2011). The theory can be helpful in predicting intentional safety violation. But it is not applicable for the assessment of errors because errors, by definition, are unintentional behaviors (Fogarty and Shaw, 2010). TPB has been applied in various fields. It has been used to predict and modify behaviors in the domain of health sciences, environmental sciences, road safety and construction safety management (Singh *et al.*, 1995; Poulter *et al.*, 2008; Jemmott, 2012; Goh *et al.*, 2018). In the construction safety research, distinctive studies have been conducted by Yang Miang Goh et al. as they have used TPB to explain the factors which influence unsafe behavior of worker (Goh *et al.*, 2018), Stephen (2005) applied TPB to predict safe behavior of workers working at height and also, Y.M. Goh & Binte Saadon (2015) have explored the factors which influence worker's behavior at height using TPB model (Goh and Binte Sa'adon, 2015). Figure 2.5 shows the TPB framework defined by Ajzen and Fishbein, (2010) where each construct has further background factors that define them.

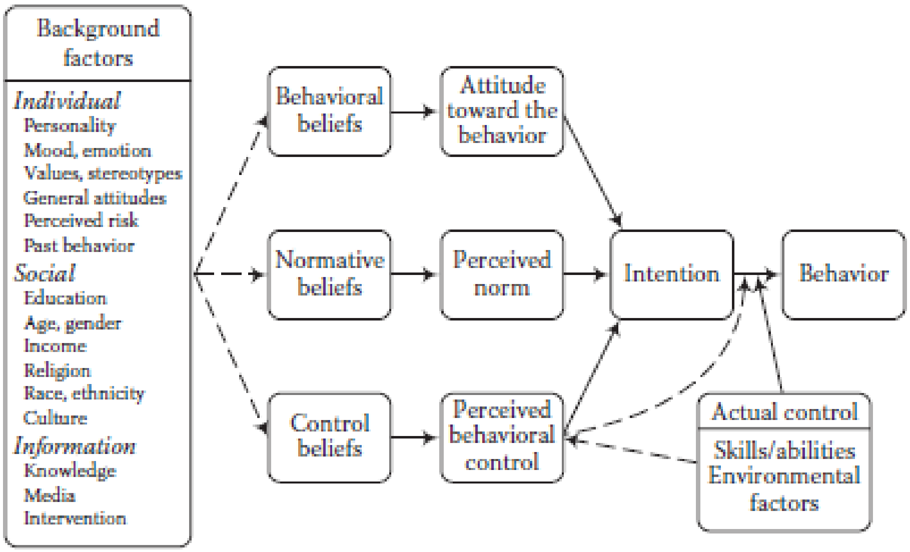

Figure 2.5: TPB framework by Ajzen & Fishbein (2010)

2.4.4. Measuring safety attitude:

There have been various studies on factors which influence behaviors but many of them are focused on either one or two variables. For example, the influence on risk perception on unsafe behaviors, influence of safety knowledge and safety participation on unsafe behavior (Guo, Yiu and González, 2016; Xia et al., 2017). The one unique aspect of this research is that it combines all the personal factors which influence safety attitude in one research instead of offering yet another fragmented research. Since factors influencing a behavior have found to have intercorrelations, thus, literature reviewed for identifying factors which influence safety attitude included studies which have reported correlations between safety attitude and other variables of unsafe behavior (Fang et al., 2004; Ismail, Doostdar and Harun, 2012)**.**

Broadly, there are two methods by which attitude towards a behavior can be measured. One way of assessing attitudes is through the Semantic Differential Scale and other is Psychometric Scales. This study has adopted the Psychometric scale as it is easy to understand. As the targeted population of questionnaire survey is Frontline workers, therefore, psychometric scale will be easy for them to comprehend.

2.4.5. Measuring perceived norms:

There is a common agreement that social environment has the ability to strongly effect actions and intentions of people. This strong impact of environment is usually described using the notion of social norm. Perceived norm is also sometimes used in place of social norms. Broadly, social norms refers to the accepted and permitted behaviors in society or in a group of people. Many theorists presume that the human behavior is directed by interest of a person and social norms restrict behavior to some extent. The major impact of the social norms in a person’s life is that his behavior besides his own interests, serves the whole social community too. Therefore, people follow the social norms as they think that the exceptions from what is normally accepted in the society can be taken as offensible and sometimes punishable too. Humans also look for purpose and meaning in social interactions. Thus, norms give meanings to their interactions by tailoring the situation and proposing guidelines about what is appropriate behavior and what is inappropriate behavior (Ajzen and Madden, 1986).

2.4.6. Measuring perceived behavioral control:

Perceived behavioral control (PBC), defined as *"the extent to which individuals think or believe that they are skilled in executing a particular behavior and have sufficient control over the successful performance of the concerned behavior".* (Terry and O'Leary, 1995; Ajzen, 2002; Kiriakidis, 2017). PBC is considered to be a predictor of behavior and has the ability to influence a behavior directly (Ajzen, 2002). For measuring PBC, direct questions about the capability of performing a task can be asked from respondents. These questions must be consistent with the construct which is under assessment. Research has revealed that questions used for assessing PBC can be categorized into two types. Perceived capacity is the perceived judgment of ease or difficulty of performing the a task and Perceived autonomy is degree of control, the person thinks he has while performing a task (Fishbein and Ajzen, 2011).

For example, the questions asked for measuring perceived capacity can be as follows:

- For me to perform behavior x would be... (very easy–very difficult).
- If I wanted to, I could easily perform behavior x… (strongly agree–strongly disagree).

Similarly, questions asked for measuring perceived autonomy are given below:

- How much control do you have over whether you perform behavior x? (no control–complete control)
- I feel in complete control over whether I perform behavior x. (completely false–completely true)

Following Table 2-2 shows the direct statements used to measure PBC in construction safety research articles.

Table 2-2: Measurement statements for PBC

Sr.	Statements for measuring PBC	References
1.	I have the necessary training(s) to work safely.	(Goh *et al.*, 2018)
2.	I have the necessary equipment to work safely.	(Goh *et al.*, 2018)
3.	I am given enough time to work safely.	(Goh *et al.*, 2018)
4.	I will work safely if I have the necessary training.	(Goh *et al.*, 2018)
5.	I will work safely if enough time was given.	(Goh *et al.*, 2018)
6.	I have complete/incomplete control over lifting materials from/to locations within my strike-zone.	(Johnson and Hall, 2005)
7.	I have complete/incomplete control over the conditions (facilities, area layout, resources, etc.) that enable me to lift materials from/to locations within my strike-zone.	(Johnson and Hall, 2005)
8.	I will fall from height if I do not hook my safety harness onto suitable anchors when working on scaffold.	(Goh and Binte Sa'adon, 2015)
9.	It is easy to hook safety harness onto suitable anchors.	(Goh and Binte Sa'adon, 2015)
10.	My coworkers and I believe that without mandatory rules, we (…) wear the PPE correctly.	(Xu, Zou and Luo, 2018)
11.	I (…) felt that it does not matter if I do not follow safety guidelines for a few times during work.	(Xu, Zou and Luo, 2018)

12.	I (…) wear PPE when I am exhausted.	(Xu, Zou and Luo, 2018)
13.	If no one reminds me, I (…) wear PPE.	(Xu, Zou and Luo, 2018)
14.	I find it easy to stick to all driving laws at all times when I am driving an LGV.	(Poulter *et al.*, 2008)
15.	The number of events outside my control which could prevent me from performing behavior x is ... (numerous–very few).	(Fishbein and Ajzen, 2011)
16.	For me to perform behavior x would be... (very easy–very difficult).	(Fishbein and Ajzen, 2011)
17.	I believe I have the ability to perform behavior x. (definitely do–definitely do not)	(Fishbein and Ajzen, 2011)
18.	I feel in complete control over whether I perform behavior x. (completely false– completely true).	(Fishbein and Ajzen, 2011)

2.5. Predictive Modeling techniques

The predictive models based upon theories used the statistical methods to explore the relationship between various behavior influencing factors. The most common methods adopted included exploratory factor analysis, confirmatory factor analysis, regression models and structural equation modelling (Fang and Wu, 2015; Johari and Jha, 2020). Although these methods provide sufficient insight into relationships between factors but they are unable to catch the complex non-linear relationships (Sato, 1995).

2.6. Machine learning (ML)

Mitchell, (2006) provided a definition of Machine Learning:“A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E”.

Machine learning is an evolving branch of computational algorithms that are designed to emulate human intelligence by learning from the surrounding environment. They are considered the working horse in the new era of the so-called big data. Techniques based on machine learning have been applied successfully in diverse fields ranging from pattern recognition, computer vision, spacecraft engineering, finance, entertainment, and computational biology to biomedical and medical applications.

2.6.1. Approaches to ML

Machine learning aproaches are classified as supervised, unsupervised and reinforcement learning. In supervised learning model, data is provided with outcomes similar to having a teacher who classifies the dataset into training examples and teaches the algorithm with the information of each example multiple times and then tests the model by giving it more data sets so that based upon its learning, it can now predict the outcome. Whereas in unsupervised learning the model can identify the outcome class information through its problem-solving without being dependent upon learnings from training datasets and reinforcement learning uses feedbacks and continouse trial and error techniques to learn from data.

Supervised learning models are further classified into two types of problems:

- **Classification** uses an algorithm to accurately assign test data into specific categories. It recognizes specific entities within the dataset and attempts to draw some conclusions on how those entities should be labeled or defined. Common classification algorithms are linear classifiers, support vector machines , decision trees, k-nearest neighbor, and random forest.

- **Regression** is used to understand the relationship between dependent and independent variables. It is commonly used to make projections, such as for sales revenue for a given business. Linera regression and logistic regression, and polynomial regression are popular regression algorithms

The use of machine learning techniques in construction safety management research has received growing interest in recent years (Liao and Perng, 2008; Goh and Chua, 2013; Goh and Binte Sa'adon, 2015; Patel and Jha, 2016; Tixier *et al.*, 2016). Past studies confirmed the high potential of the machine learning algorithms in evaluating data related to safety management. For example, Arciszewski et al., (1995) used a machine learning technique to transform accident reports into decision rules for better safety practices. Similarly, Tixier et al., (2016) used injury reports as inputs and applied Stochastic Gradient Tree Boosting and Random Forest techniques for the prediction of the outcomes including injury type and severity and the injured body part. Goh and Chua, (2013) worked on the severity of injuries by using audit scores of safety management systems for the training of ANN model.

2.6.2. ML for behaviour prediction

One goal of machine learning is modelling the human mechanisms related to human learning methods. In the psychological framework, reserachers have developed learning algorithms that are consistent with knowledge of the human cognitive architecture and are designed to explain specific observed learning behaviors. Behavior prediction domain has a very wide range where both the human behavior and the mechanism of the predictive models are called the system behaviours. Artificial nueral nets and deep neural networks are most widely used ML techniques for human behavior prediction. ANN is used in research of Patel and Jha, (2015b, 2015a) and (Goh and Binte Sa'adon, 2015) for the prediction of unsafe behvaiors and safety culture variables. Goh and Binte Sa'adon, (2015) also used DT and reported that both ANN and DT provide better prediction results as compared to linear regression. Yang *et al.*, (2016) applied SVM with a pattern-recognition algorithm for the detection of near-miss incidents of ironworkers and achieved 87.5% accuracy in predictions. In addition, Akhavian and Behzadan (2016), used five machine learning methods for recognition and classification of activities of construction workers. The results showed that the neural networks outpaced other ML classifiers, having an accuracy of 97%.

2.7. Artificial Neural Networks (ANN)

Artificial neural networks (ANNs) are a subset of ML or can be called as a supervised ML technique. ANNs have the ability to learn from data and update themselves (Networks *et al.*, 1994; Boussabaine, 1996). In simpler words, an ANN model is based on human neural biology and is defined as an information-processing network, developed through the generality of mathematical simulations. These systems do not need any programming or specific set of instructions to work (E. Alpaydin, 2020). The concept of ANNs came from the structure of human brain where numerous neurons connect and process information. Their basic principle is that they learn from past experiences, simplify the lessons learnt from previous examples, extract information and characteristics of similar problems, then based upon all of these, an output is generated (Boussabaine, 1996). Network elements interact together in a way which is similar to neural connections in human nervous system and allow signals to travel in parallel and series through the network (Rosa *et al.*, 2020). The input data is first split into a 70/30 ratio, where 70% of data is used for training of model. During the training, an ANN model needs data from both the inputs and outputs of a given problem and trains itself in a way that if new inputs are given to the system, ANN can predict the outcome based upon its previous learning from the data. The model uses the other 30% data to validate its learning by constantly testing itself. Notable results have been achieved through various ANN models developed in the past decades. Out of all these models, the non-linear multi-layered networks have found their applications in many fields (Zupan, 1994; Weber et al., 1999; Tam and Tong, 2003; Rebaño-Edwards, 2007; Patel and Jha, 2016; Polat et al., 2016; Peško et al., 2017; Zhang, Cao and Zhao, 2019).

2.7.1. Elements of ANN

An ANN consists of nodes which are connected by supervised links. An input layer, hidden layers and output layer form the basic configuration of an ANN system with directed links between the nodes of each layer. Figure 2.6 shows general architecture of ANN with weights on neurons in input, hidden and output layers. A numeric weight is assigned to each of the link.

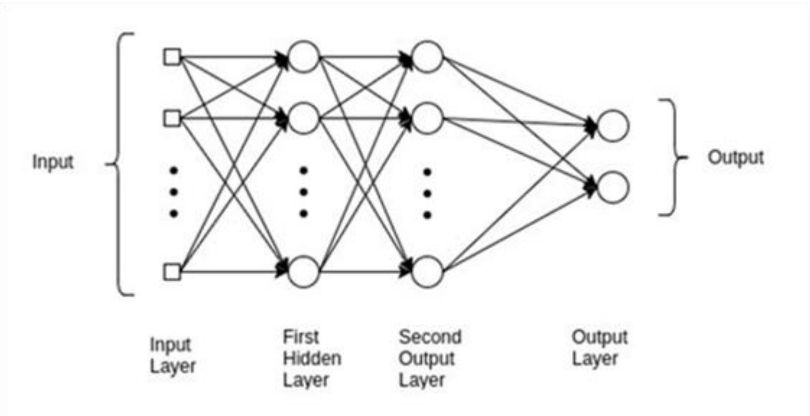

Figure 2.6: ANN structure

The number of hidden layers in an ANN model can be adjusted according to the error and trial method. Most of the times, one or two hidden layers are enough for a system to predict the output effectively. The other components of ANN include the activation functions, loss function and accuracy metrics, optimizer and a propagation method. Hidden layers have an activation function which processes information after applying a mathematical generalization (Networks *et al.*, 1994). Loss function are the objective functions that measure the total loss occurred in predicting the correct class of dataset. Accuracy measurements include the number of times the model was successful in correctly predicting the outcome. The propagation method defines the main algorithm of model and provides a method on how the neurons will update themselves during the learning process.

These parameters play an important rule in decreasing the loss and getting a high accuracy. Figure 2.7 displays the weights and connections between layers.

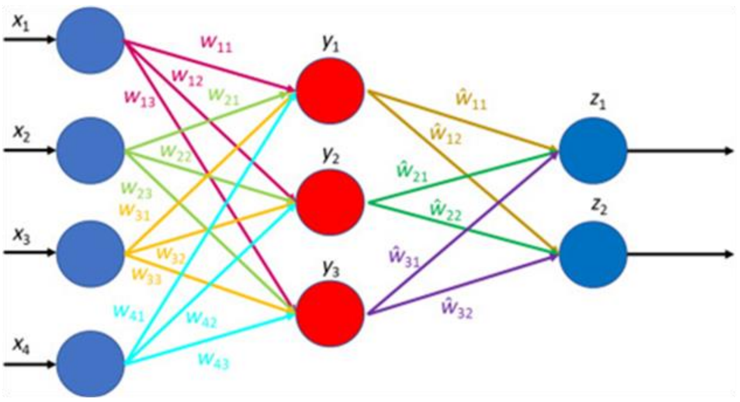

Figure 2.7: ANN weights and links

The optimizers use the data history and momentum to keep track of the updates made to weights each time the model goes through forward and backwards propagations until the desired accuracy and loss values are achieved (E. Alpaydin, 2020). The number of times a model is run is called epoch. Figure 2.8 taken from Boussabaine (1996) shows the forward and backward propagation through neurons with loss values. There are other hyperparameters involved in the background computations such as bias, momentums, learning rates and dropouts. There can be problems of overfitting and underfitting of data. If a network has more nodes and layers than the number of training sets, data is overfitted. Such overfitted data result in a very low training error, which is not a good performance measuring criterion.

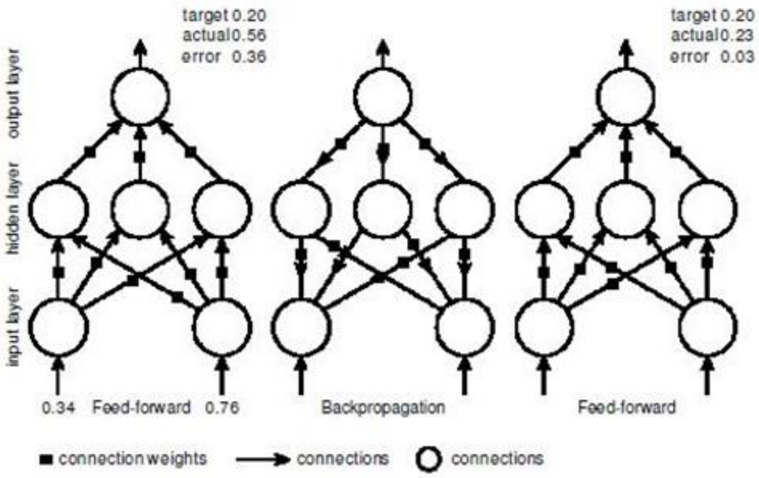

Figure 2.8: Feedforward & backward propagation of ANN

On the other hand, a smaller number of input nodes and hidden nodes result in underfitting, which further results in poor performance of the model (Alwosheel, van Cranenburgh and Chorus, 2018). To overcome the problem of overfitting, the dropout technique is used, which considers a specific number of neurons during each propagation and drops the left-out neurons. In the next epoch, the left-out neurons are used, and previously active neurons are kept as dead (Boluki *et al.*, 2020). Figure 2.9 show a best-fitted model as described by Alwosheel, van Cranenburgh and Chorus, (2018).

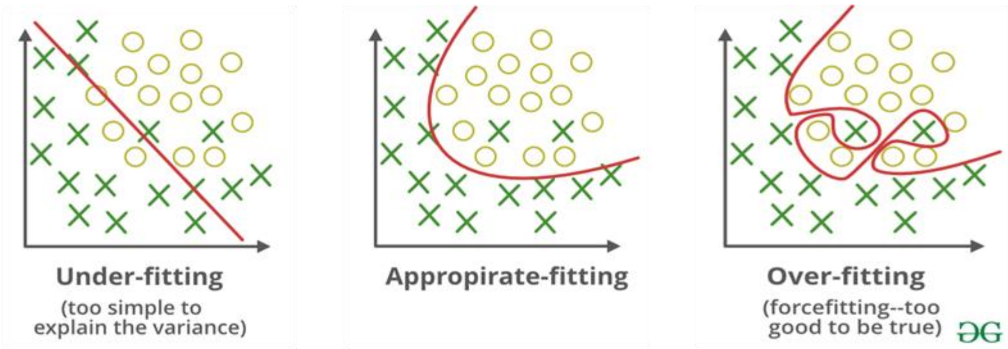

Figure 2.9: ANN's fitness criteria

There are also various approaches to ANN algorithms. One other approach is k-cross-folds techniques. This approach is usually adopted when dataset has less samples and accuracy values may not be reliable when using a data split of 70,30 for training and validation.

2.7.2. Advantages of ANN

ANNs possess distinct features which offer many advantages as compared to the traditional modeling methods. ANNs have been very successful in determining complicated and unknown relationships present between variables of a dataset. ANNs are self-adaptive, and are data driven as they have the capacity to catch complex functional connections in the data set. They optimize their output by automatically adjusting their weights unlike the statistical models (Boussabaine, 1996). They are also able to accurately deduce complicated nonlinear relationships as compared to the statistical models e.g., regression models (Networks et al., 1994; Bento, Cardoso and Dias, 2005; Rosa et al., 2020). ANNs are being used for classification, vector quantification, pattern association, function approximation and prediction.

2.7.3. ANN for behavior prediction (psychology)

Since the development of neural networks, evidence is that these ANN models have a close proximity with the field of psychology. Although artificial networks are used more frequently in fields as financial analysis, economical modelling and marketing studies, their use in areas of psychology has provided lots of hopeful discoveries. Artificial neural networks have been used successfully in the prediction and diagnosis description of various psychiatric disorders such as eating disorders, depression, compulsions, or schizophrenia. In conclusion, artificial neural nets offer a promising alternative of research approach for modern-day psychiatry and clinical psychology (Wu and Feng, 2018). Whereas limited work has been done in the domain of human behavior prediction. Recently, using ANN, a predictive model for assessing violent behaviors in human was proposed using various stress measuring scales. (Ramón and Rodríguez, 2020). A study conducted by Abubakar et al., (2019) used ANN model to analyse and predict knowledge hiding behaviour of humans (Abubakar *et al.*, 2019). Another study has used hand writings to predict human behaviour using ANN model (Champa and AnandaKumar, 2010). The recurrent neural networks have also been used to predict human bahaviour through human actions and activities and allows to predict next step of user (Almeida and Azkune, 2018).

2.7.4. ANN adoption in Construction

In construction management, ANNs have been used widely, mostly in the sub-domains of cost prediction (Weber *et al.*, 1999; Emsley *et al.*, 2002; Attalla and Hegazy, 2003; Polat, 2007; Cheng, Tsai and Liu, 2009; Jha and Chockalingam, 2011; Polat *et al.*, 2016; Peško *et al.*, 2017; Juszczyk, Leśniak and Zima, 2018; Abd and Naseef, 2019; El-Kholy, 2019; Sandhya and Philominal Judit, 2020; El-Kholy, Tahwia and Elsayed, 2020) and safety management (Wei and Lee, 2007; Ciarapica and Giacchetta, 2009; Goh and Binte Sa'adon, 2015; Patel and Jha, 2015a, 2015b, 2016; Tixier *et al.*, 2016; Goh *et al.*, 2018; Zhang, Cao and Zhao, 2019; Bangaru *et al.*, 2021). ANNs are also adopted in contract management (Al-Sobiei, Arditi and Polat, 2005; Chen and Hsu, 2007) and for the prediction of construction project performance (Chao and Skibniewski, 1994; Cheung *et al.*, 2006; Rebaño-Edwards, 2007; El-Gohary, Aziz and Abdel-Khalek, 2017). Figure 2.10 shows the adaptation of ANNs in various subdomains of construction management and shows that construction safety is the sub-domain where ANN has been most frequently used.

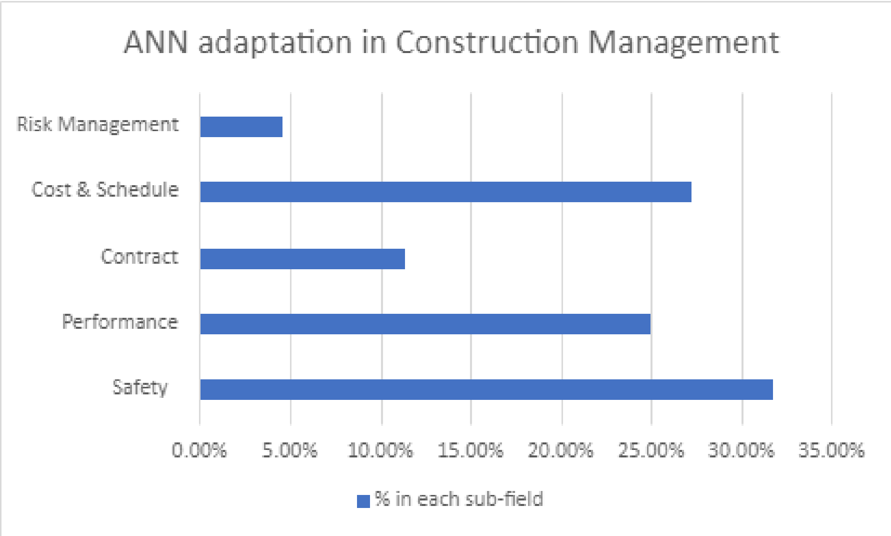

Figure 2.10: Use of ANN in construction management's sub-domains

2.7.5. Software Packages for ANN

The ANNs are usually implemented by using various computer software packages which are highly advanced and readily available for the training of ANNs (McDermott, 2009). These packages are TensorFlow, MATLAB neural network toolbox, Neuro-Solutions, and the Alyuda Neuro-Intelligence. There are also separate libararies specially designed for neural networks such as Keras API and Scikit

Learning. These software packages help in solving many real-world problems by assisting in designing, training, testing and validation of thousands neural network models. The processes in such software packages are automated, providing results in the form of graphs and report for easily understanding of underlying behavior and working of the phenomena under consideration.

2.8. Chapter conclusion

This chapter included literature review on a wide range of construction safety management research topics such as causes of accidents in construction, unsafe behaviors of workers, methods and theories to analyse and evaluate these unsafe behaviors and behavior modelling techniques. The discussed literature will be adopted to develop a unsafe propensity prediction model so that accidents caused by worker's behavior can be monitored and minimized by the construction organization.

Chapter 3

Research methodology

Overview:

This chapter includes the methodological framework which is used to conduct this research. The research work starts with the identification of a research gap and then moves forward to formulate the research objectives. After that, a literature review is carried out to get insight of what has been done earlier in this sub-domain of construction management. This chapter explains each step of the research in detail.

3.1. Research Methodology Framework

The research methodology as shown in Figure 3.1 is conducted in four phases which are explained below.

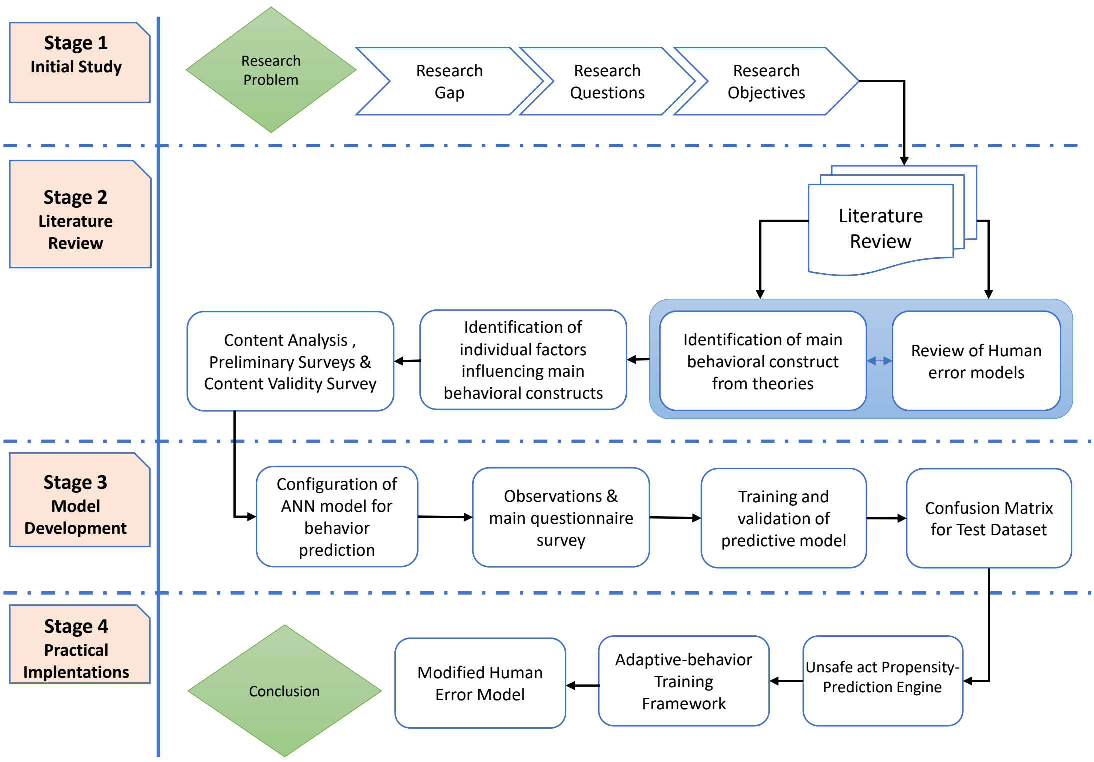

Figure 3.1: Research methodology framework

3.1.1. Initial Study

The first stage of this study involves the identification of research gap after the exploration of research problem. Research questions are used to get a direction for the study and then finally research objectives are formulated. The research gap is identified through an extensive study of available literature in journal articles, conference papers and books on safety management, human error models and theories of behavior prediction and modification. The research objectives of this study are formulated after careful consideration of all the trends and literature available in the field of construction safety management.

3.1.2. Literature review

The second stage of the study involves a detailed literature review of all the concepts and techniques used in this study.

3.1.2.1. Review of human error models

The literature review started with the study of human error models since this research is concerned with individual unsafe acts and their causes. It was noticed that in the earlier literature studies unsafe behaviors are most common cause of accidents as unsafe behaviors lead to unsafe acts, which lead to an accident. The data has been acquired from various research articles on accident causation models and causes of unsafe behaviors are also explored through literature review.

3.1.2.2. Identification of behavioral constructs from TPB

The methodology implemented in this study incorporates theoretical constructs, taken from the field of psychology. The consrtructs of TPB have been utilized. The constructs selected are Attitudes and Perceived behavioral control (PBC).

3.1.2.3. Identification of individual factors

The selected behavioral constructs i.e., attitude and PBC are further influenced by the background factors. Most of the past research studies focus on the organizational factors which influence behavior of workers, and a very minute amount of work has been done on personal factors which influence workers behavior. Since an organization can only ensure a strong safety culture if it provides no accommodation for unsafe behaviors leading to individual unsafe acts. Furthermore, the articles on personal factors of workers have targeted one or two specific factors and thus the literature available is very much fragmented. Thus, the factors identified from literature are given a frequency score and literature score.

3.1.2.4. Content analysis and field surveys

An expert opinion survey was conducted to get the view from the construction safety managers, behavioral scientists, and psychologists on the background individual factors affecting attitude of workers. A field score is calculated and final score for each factor is calculated through 60field-40literature impact ratio. A content validity survey was conducted to get validate the measurement statements that were used to get the reponses of workers in the main survey. Furthermore, to cross-validate the shortlisted factors, a field survey was carried out, where semi-structured interviews of workers were conducted. The interviews were analyzed to extract the factors that influence their behaviors.

3.1.3. Model Development

The third stage of the research comprises of the development process of an ANN based prediction model. The steps carried out to make the model as explained below.

3.1.3.1. Configuration of ANN model

The first step of this stage includes the formulation of an ANN architecture at first since performance of an ANN model highly depend on the architecture of network. It includes the initial selection of the number of layers, neurons and mathematical functions.

3.1.3.2. Observations & questionnaire

The data is collected through observations and questionnaire. This step provided the datasets with inputs and outputs for the learning of ANN model.

3.1.3.3. Training and validation of model

The model is executed on the Tensorflow for its training that comprises of learning from dataset. Model simultaneously checks its training accuracy and validation accuracy during each epoch. It is run with varying number of ANN elements, also called as hit and trial method, to get the best parameters of the model.

3.1.3.4. Confusion metrics for test dataset

The test data set comprised of 10% of the total datset and it is used to test the model's performance. The model is not given the output class of the sample and is required to predict the target. The confusion metrics allows the visual representation of the number of times the model successfully predicted the actual class.

3.1.4. Practical implementations

The third stage of the research consists of implementations of the developed model. It included a web-based engine associated with a training framework. It also gives a modified HFACS model.

3.1.4.1. Unsafe act propensity prediction engine

The propensity prediction engine is a web-based tool with a user interface, powered with the developed ANN model to predict the propensity for unsafe acts. This engine can be used by construction safety managers to classify workers into categories on the basis of their behaviors.

3.1.4.2. Adaptive-training framework

Based upon the prediction of the ANN engine, the workers can be given specialized trainings and behavior interventions so that their unsafe behaviors can be tackled and accidents can be avoided.

3.1.4.3. Modified HFACS

The final step is the modification in human error model with the suggestion of adding the behavioral precursor under the preconditions of unsafe acts defence level.

3.2. Chapter conclusion

This chapter consists of the methodology adopted by the researcher to achieve the research objectives. It contained a methodological framework and explanation of each stage and sub-stage of the research.

Chapter 4

Analysis and Results

Overview:

This first part of this chapter consists of explored individual factors through literature and field surveys. It also enlists the justifications behind using specific behavioral constructs. The second part consists of detailed data collection method while the third part comprises of ANN model development and execution.

4.1. Selection of behavioral constructs from TPB

Since this research is focused on predicting proness for individual safety acts caused by personal preferences of workers. The study is keeping the construct of perceived norm as a constant. This is due to the fact that perceived norms are influenced by group norms and organization's safety culture (Ajzen and Fishbein, 2010). If we are to see if a worker is involved in an intentional unsafe act, it is important to assume that factors outside his volitional control are constant. i.e., he has been given necessary training, PPE's and no unsafe condition was present. Since according to Ajzen and Fishbein (2010), the constructs of the theory can be modified or used as per the context of the problem, therefore, this study takes the construct of attitude and perceived behavioral control (PBC) only, for predicting unsafe acts of workers. Intentions partially mediate the relationship between Attitude and Behavior and can influence the safety behavior of workers (Fogarty and Shaw, 2010). But, since we are interested in catering propensity for intentional unsafe acts of workers, the effect of intentions is ignored. Figure 4.1 shows the TPB-based theoretical foundation developed for this study.

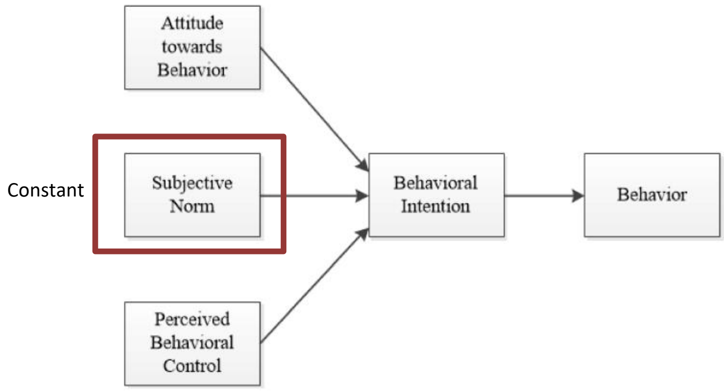

Figure 4.1: TPB-Based individual unsafe behavior model

4.1.1. Individual factors for behavioral constructs

The inputs variables are the individual factors which effect a person's behavior. Since, this study has adopted the TPB to explain the variances in individual behaviors, the background features for each of construct are explored through literature review first.

4.1.1.1. Safety Attitude (literature):

A total of 19 factors for safety attitude are identified. The identified factors are named as features of safety attitude and are given a literature score on a three-point Likert scale (1=Low, 3=Medium and 5=High), based on frequency of its occurrence in literature and its significance, as assessed by each respective author. Henceforth, the literature score was calculated for each factor by finding the product of its frequency and impact score, respectively. The literature score was also normalized before using it for further analysis. Afterwards, the identified factors of safety attitude are ranked.

Table 4-1 includes the personal factors/ variables which influence safety attitude of workers, along with the references and literature score.

Table 4-1: Individual factors influencing behavior

Sr.	Factor	Source	Literature Score
1	Perceived risk	(Hofmann and Stetzer, 1996; Tomás, Meliá and Oliver, 1999; Mullen, 2004; Choudhry and Fang, 2008; Mohamed, Ali and Tam, 2009; Fugas, Silva and Meliá, 2012; Chi, Han and Kim, 2013; Goh and Chua, 2013; Fang, Zhao and Zhang, 2016a; Xu, Zou and Luo, 2018; Guo, Goh and Le Xin Wong, 2018)	**0.384**
2	Safety knowledge	(Hofmann and Stetzer, 1996; Hollnagel, 2002; Fang *et al.*, 2004; Mitropoulos, Abdelhamid and Howell, 2005; Choudhry and Fang, 2008; Mohamed, Ali and Tam, 2009; Wu, Gibb and Li, 2010; Zhou, Fang and Mohamed, 2011; Ismail, Doostdar and Harun, 2012; Yu *et al.*, 2014; Shin, Gwak and Lee, 2015; Guo, Yiu and González, 2016; Wang, Zou and Li, 2016)	**0.300**
3	Safety participation	(Tomás, Meliá and Oliver, 1999; Fang *et al.*, 2004; Zhou, Fang and Mohamed, 2011; Yu *et al.*, 2014; Xu, Zhang and Hou, 2019)	**0.138**

4	Work environment uncertainty	(Choudhry and Fang, 2008; Mohamed, Ali and Tam, 2009; Cheng, Lin and Leu, 2010; Hollnagel, 2013; Yu *et al.*, 2014 ; Shin, Gwak and Lee, 2015)	**0.161**
5	Work experience	(Tam and Tong, 2003; Aksorn and Hadikusumo, 2007; Zhou, Fang and Wang, 2008; Choudhry and Fang, 2008; Liao and Perng, 2008; Mitropoulos, Cupido and Namboodiri, 2009; Mohamed, Ali and Tam, 2009; Cheng, Lin and Leu, 2010; Ismail, Doostdar and Harun, 2012; Hollnagel, 2013)	**0.230**
6	Overconfidence	(Mullen, 2004; Aksorn and Hadikusumo, 2007)	**0.046**
7	Hazard awareness	(Hofmann and Stetzer, 1996; Choudhry and Fang, 2008; Cheng, Lin and Leu, 2010; Fugas, Silva and Meliá, 2012; Yu *et al.*, 2014)	**0.192**
8	Macho syndrome	(Mullen, 2004; Aksorn and Hadikusumo, 2007; Choudhry and Fang, 2008)	**0.069**
9	Competence	(Mullen, 2004; Choudhry and Fang, 2008)	**0.046**
10	Avoiding teasing	(Mullen, 2004; Choudhry, Fang and Ahmed, 2008)	**0.046**

11	Age	(Liao and Perng, 2008; Cheng, Lin and Leu, 2010; Chen and Jin, 2013; Wang, Zou and Li, 2016)	**0.092**
12	Communication	(Ismail, Doostdar and Harun, 2012; Shin *et al.*, 2014; Yu *et al.*, 2014)	**0.069**
13	No. of dependents	(Wang, Zou and Li, 2016)	**0.023**
14	Job security	(Shin, Gwak and Lee, 2015; Wang, Zou and Li, 2016)	**0.046**
15	Safety motivation	(Hofmann and Stetzer, 1996; Hollnagel, 2002; Wu, Gibb and Li, 2010; Zhou, Fang and Mohamed, 2011; Guo, Yiu and González, 2016)	**0.038**
16	Perceived barriers	(Choudhry and Fang, 2008; Zhou, Fang and Mohamed, 2011; Chen and Jin, 2013)	**0.153**
17	Beliefs	(Hofmann and Stetzer, 1996; Brown, Willis and Prussia, 2000; Choudhry and Fang, 2008; Mohamed, Ali and Tam, 2009; Chen and Jin, 2013; Fang, Zhao and Zhang, 2016a; Xu, Zhang and Hou, 2019)	**0.307**
18	Accident Experience	(Fang *et al.*, 2004)	**0.038**

19	Awareness of utility of outcome	(Fang *et al.*, 2004; Mitropoulos, Cupido and Namboodiri, 2009; Chen and Jin, 2013; Shin *et al.*, 2014; Guo, Yiu and González, 2015)	**0.115**

4.1.1.2. Preliminary questionnaire for safety attitude factors

After literature analysis, a preliminary survey was performed to include input from field professionals as well, for the purpose of ranking these factors. A preliminary survey questionnaire was drafted and then circulated to experts of all disciplines being involved in this study, from both developed and developing countries. The reason for targeting both developed and developing countries is that this study considers the construction organization's safety culture, whether that organization be in a developing country or developed country. Also, the social construct of TPB was kept as a constant too since the impact of social norms of a society on a worker is not a personal factor. Total 30 responses are collected. A Cronbach alpha test for the measuring the internal consistency of the questionnaire data is also adopted. The value of Cronbach's alpha came out to be 0.80 which is within an acceptable range. The Table 4-2 shows the information of the respondents of preliminary survey. "C" represents the category and "F" is frequency that shows the number of respondents of each category.

Table 4-2: Information of respondents of preliminary survey

Preliminary Survey; Respondent's Demographics							
Level of Education		Years of experience		Field of Study		Country of work	
C	**F**	**C**	**F**	**C**	**F**	**C**	**F**

16 years	4	0-5	6	Sociology	1	Pakistan	16
18 years	9	6-10	17	Psychology	7	China	5
Phd	17	11-20	7	Construction Safety Management	17	UK	3
		>21	0	Environmental Health & Safety	2	India	2
				Civil Engineering	3	Thailand	1
						Korea	1
						Australia	1
						NewZealand	1
Total			**30**				

From the preliminary survey, the field score was also calculated and then normalized. Based on 60% field and 40% literature impact ratios, 8 factors out of 19 are selected on simple majority principle having 60% cumulative impact. Table 4-3 enlist the shortlisted factors for safety attitude measurement.

Table 4-3: Selected factors of safety attitude

Code	Factor	C.Score
SA-1	Safety Knowledge	0.114
SA-2	Perceived Risk	0.212
SA-3	Work Experience	0.305
SA-4	Beliefs	0.392
SA-5	Hazard Awareness	0.463

SA-6	Work Environment uncertainity	0.518
SA-7	Age	0.603
SA-8	Perceived Barrier to safety	0.6500

4.1.1.3. Measurement items for safety attitude

Previously, the reserachers of construction safety have used statements as measurement items for attitude factors (Johnson and Hall, 2005; Poulter *et al.*, 2008; Fogarty and Shaw, 2010; Goh and Binte Sa'adon, 2015; Fang, Zhao and Zhang, 2016b; Goh *et al.*, 2018; Xu, Zou and Luo, 2018; Chen and Yang, 2019). This study has modified them and verified from psychological experts to see if the items actually measure the underlined factor or not. Same is done for PBC factors, except that only two factors i.e., Autonomy and Capacity are taken from literature review (Fishbein and Ajzen, 2010). Measurement items for factors of PBC are also adopted from previous studies by (Hofmann and Stetzer, 1996; Johnson and Hall, 2005; Fogarty and Shaw, 2010; Goh and Binte Sa'adon, 2015; Fang, Zhao and Zhang, 2016b; Goh *et al.*, 2018; Xu, Zou and Luo, 2018) and then are also validated. Table contains the measurement items for individual factors of safety attitude.

4.1.1.4. Factors for PBC

For the measure of Perceived Behavioral Control, direct statements on easy-difficult scale are used (Fishbein and Ajzen, 2011). These statements are also explored through intensive literature review of articles and books on behavioral modification. This study has used direct questions from previous studies and modified them for measuring PBC and has incorporated both categories of PBC i.e., perceived capacity & perceived autonomy. Table 4-4 contains the modified measurement items for perceived capacity and autonomy factor. The code in table corresponds to the each construct and its item number. SA-1 describes the first factor for safety attitude. Item represents the measurement statements as some of the factors have been given two measurement statements.

Table 4-4: Measurement items for SA and PBC factors

Construct 1			**Safety Attitude**
Code	**Factor**	**Item**	**Measurement statements**
SA-1	Perceived risk	A	I think the likelihood of falling from height and getting injured is more.
		B	I think the possible consequences of fall accidents are severe.
SA-2	Work experience	A	From my own work experience on construction sites, I think safety procedures implementation is important.
SA-3	Safety knowledge	A	I am aware of risks associated with my job and health safety.
		B	I am aware of necessary provisions to be taken while performing the job at construction sites.
SA-4	Work environment uncertainty	A	I think Personal protective equipment (PPEs) is effective for our safety at construction sites.
		B	I think avoiding hazards in our work environment is possible.
SA-5	Hazard awareness	A	I can identify potentially hazardous situations.

SA-6	Age	A	I believe that anyone can be involved in an accident either young or elder.
		B	I believe that everyone on the worksite either young or elder should comply with safety rules.
SA-7	Beliefs	A	I believe that an accident should not be regarded as an act of nature as it is possible to avoid accidents by complying with safety rules.
		B	I believe that my safety is my own responsibility.
SA-8	Perceived barriers	A	I believe that most of the safety procedures are convenient.
		B	I believe that all rules and policies relevant to my job are practical.
	Construct 2		**Perceived behavioral control**
PBC-1	Capacity measurement	A	For me to work at height without any PPE would be... (easy/difficult).
		B	I believe that I can…… perform my work without having an accident. (definitely/definitely do not).
PBC-2	Autonomy measurement	A	I have …. Over my task without a PPE. (Complete control- no control).
		B	I believe that the number of external events outside my control which can cause an accident while doing work at height is.... (none/numerous)

4.2. Content validity analysis of measurement items:

This study conducted a content validity survey to validate the statements that are used to measure SA and PBC. The content validity survey includes five (5) professional experts from the field of psychology and social sciences. They are requested to rate every one of the measurement items in terms of their application to the primary construct using a four-point scale (1-not relevant, 2-somewhat relevant, 3-quite relevant, and 4-highly relevant (Polit and Beck, 2006). Item-level content validity index is determined by the number of experts providing a score of either three or four; divided by the overall number of experts. Items with an item-content validity index of 1.0 are kept (Polit and Beck, 2006). Table 4-5 contains the finalized factors, their code names, and their respective measurement items validated through content validity survey. The SA responds to "safety attitude" whereas each expert giving ratings to each measurement item on a 4 point Likert scale is reported under E-1 to E-5. E-1 to E-5 are the 5 experts. Item A and B are the modified measurement items of individual factors. Since the item-level content validity of the items came out to be 1, thus, they were retained.

Table 4-5: Item validity index

Content Validity Assessment						
Scale Items	**Expert ratings (5 experts) on a 4-point scale**					**Item-validity index**
SA	E-1	E-2	E-3	E-4	E-5	
1	Safety Knowledge					
Item A	3	4	4	4	4	1
Item B	3	4	4	4	4	1
2	Perceived Risk					
Item A	3	4	4	4	4	1

Item B	3	4	4	4	3	1
	3 Work Experience					
Item A	3	4	4	4	3	1
	4 Beliefs					
Item A	3	3	4	4	4	1
Item B	3	4	4	3	4	1
	5 Hazard Awareness					
Item A	3	4	4	4	4	1
	6 Work Environment Uncertainity					
Item A	3	3	3	3	3	1
Item B	3	4	3	3	4	1
	7 Age					
Item A	3	4	4	3	4	1
Item B	3	4	4	3	4	1
	8 Perceived Barrriers to safety					
Item A	3	4	4	3	3	1
Item B	3	4	3	4	3	1
PBC	**E-1**	**E-2**	**E-3**	**E-4**	**E-5**	**Item-validity index**
	9 Capacity Measure					
Item A	3	4	4	4	3	1
Item B	3	4	4	3	3	1
	10 Autonomy Measure					
Item A	3	4	4	3	3	1

Item B	3	4	4	4	3	1

4.3. Cross-validity of Observed factors

A field survey was conducted at a local construction site to cross-check whether the factors shortlisted for individual behavior prediction are valid in prespective of developing countries. Seven (7) workers from multiple trades are interviewed via semi-structured questionnaire. An interview guide is developed with 8 items that were to be used to steer the interview in case the worker deviates from the topic under consideration. The questions included seven (7) questions regarding worker's experiences, age, safety concerns, safety trainings, accident causations, current safety measures of their organization and use of PPE's. All questions are simplified in order to get what workers think of safety masures in general. The constractor had provided the worker's with necessary PPE's and site supervisors occassionlly check worker's compliance with PPE's. These interview items are open-ended questions and were designed in a way that respondents do not feel that they are being questioned about their safety performance. This is due to the fact that construction workers are aware of their unsafe acts and if they are questioned about safety, they become hesitant and do not answer openly. They also tend to think that interview is some kind of evaluation and telling the truth might result in losing the job. Therefore, each interview was conducted in the absence of any site official. Workers were informed that the whole activity is solely for the research purpose. A total of 7 interviews were conducted when saturation is observed. Table contains the personal information of the respondents. Furthermore, respondents were also uncomfortable with recording devices present and managers asked for permission to record their interviews. None of the 7 workers were permitted to record their interviews. This also shows the mistrust of workers on safety managers as they think that their views might be exploited against them in some way.

4.3.1. Semi-structured interviews

The interviews were conducted in the native language, as most of the respondents have received primary-level education. Since the audio recording was denied by workers, the research relied on interview scripts and notes. To ensure the validity of the research, previous literature on the validity of interview scripts was explored. According to Nordstrom (2015), the presence of a recording device even if it's turned off influences the interactions of respondents by making them nervous about the situation. It is further suggested that the accuracy of data can be misleading as recording focuses the attention on what are the words of respondents, while the focus of research should be on personal interactions and observations too. Also, it is assumed that the information taken from interviews cannot be labeled as accurate since the respondents may give false information to shield their privacy or personal opinions and it is very much possible that the respondents tell what the interviewer wants to hear, instead of his actual thoughts. Thus, Rutakumwa et al., (2020) concluded that a choice for interview method should be given when the measurement questions are sensitive to the respondents and the presence of a recording device can influence the accuracy of the information. This study, thus, opted for interview scripts and notes. Since taking notes during interviews can be a challenging task and the interviewer may skip an important piece of information. To overcome this weakness, the participants were not given any time window, they were requested to contribute to research by giving their opinions on safety on construction sites either in the form of one sentence or a whole story. Table 4-6 contains the age, work experience, education and work type of interviewees.

Table 4-6: Information of respondents of cross-validity survey

Sr.	Respondent's Age (yrs.)	Work Experience (yrs.)	Education Level	Work trade & type
1	34	8	Secondary	Scaffolder
2	30	9	Primary	Concreter
3	28	13	Uneducated	Steel binder
4	48	21	Primary	Mason
5	37	18	Primary	Concreter
6	25	6	Secondary	Concreter
7	33	7	Secondary	Steel binder

4.3.2. Analysis of scripted notes

The interview scripts were analyzed through a grounded theory approach. The grounded theory is a qualitative analysis method that provides a systematic and organized way of analyzing the data through careful consideration of conversations (Locke, 2002). The grounded theory approach specifically is an inductive method to qualitative research that focuses on creating theory from the collected data (Williams and Moser, 2019). In this approach, the data collected through interviews or observations demands the researcher to acknowledge the thematic connectivity and emergence of facts as the whole process including data analysis and interpretation of results involves a constant interaction between the data and researchers (Khan, 2014). There are three coding steps involved in applying the grounded theory: open, axial, and selective coding. Coding is a cyclic process that revolves around the concept of finding the underlying perceptions of subjects and seeking the dimensions of the research topic under study. Each sentence in each script note was analyzed systematically through repeated re-readings to understand the similarities and contradictions across the interviews. The common pieces in conversations of respondents are identified and named as open codes. The sentence fragments are given a name that corresponds to an underlying factor. For example, “sometimes free movement is restricted due to PPE” is named as “perceived inefficiencies of PPE”.

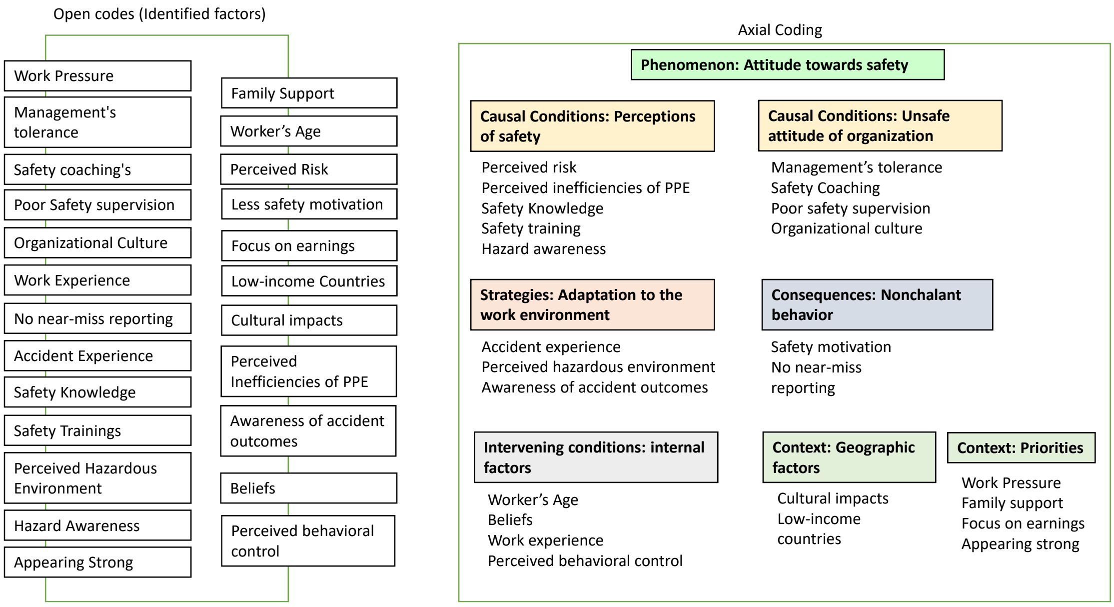

Figure 4.2: Open codes to axial codes

Axial coding further aligns, refines, and categorizes the identified factors. It also gives an insight into the relationships between factors (open codes), so that selective coding can be done. For the sub-categories in axial coding, usually, the categories named as a phenomenon, causal conditions, strategies, consequences of phenomenon, context, and intervening conditions are constructed (Williams and Moser, 2019). Causal conditins are the factors that cause the phenomenon. From the Figure 4.2 the causal conditions included the factors related to perceptions of safety and unsafe attitude of organization towards safety management. Strategies are the potential actions and responses of the subjects in order to adjust to an environment and consequences are the outcomes of the strategies. Context can be defined as the circumstances where strategies are set whereas intervening conditions responds to the factors that mediate the relationships between a phenomenon and its causal conditions (Williams and Moser, 2019). The factors grouped under these axial sub-categories are also listed in Figure 4.2.Selective coding is the last level of coding. It involves the selection and integration of sub-categories of from axial coding into a meaning-filled expression leading to the formulation of theme of the study (Williams and Moser, 2019). The theme formed is named as "critical factors affecting safety attitude of workers". Figure 4.3 further explains the sub-categories and theme in a confined way.

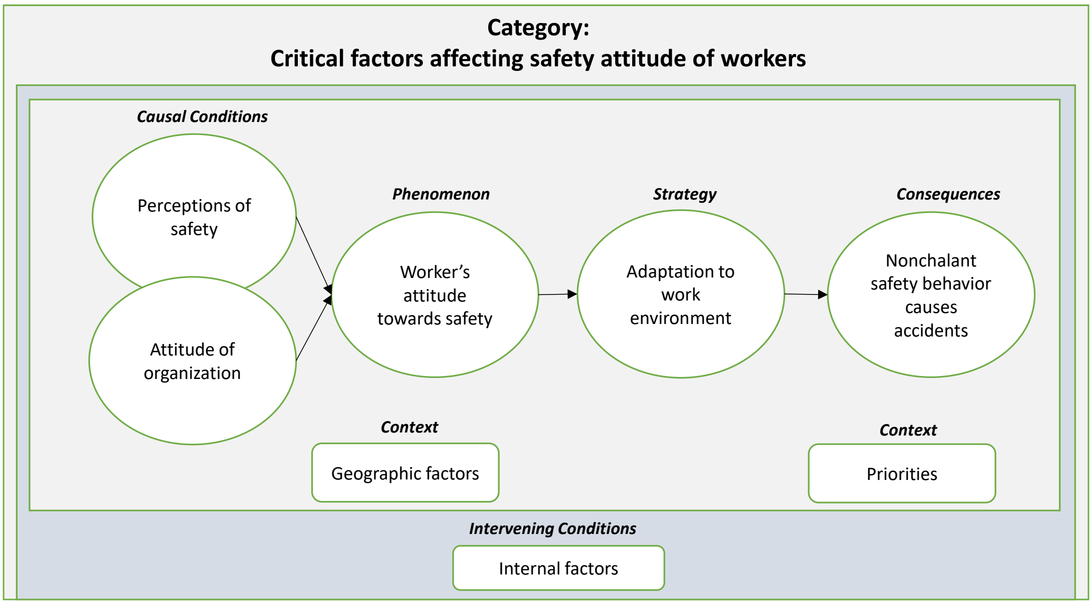

Figure 4.3: Constructing category from axial coding

Thus, from this analysis, it was validated that the factors taken from theory are almost similar to the ones reported by workers, thereby confirming the theoretical foundation constructed for individual unsafe acts of workers.

4.4. Initial configuration of ANN Model

After the finalization of inputs, an initial architecture for the ANN model was developed. Initially, a multi layer feed forward model was configured with a total of 8 neurons in input layer, 9 neurons in 1st hidden layer, and an outpout layer with 2 neurons. i.e., for two classes. The input layers are the variables of Safety attitude coded as SA-1 to SA-6 according to their ranking obtained from the content analysis. Two inputs of PBC are also added i.e., the capacity and autonomy factors. These are coded as PBC-1 and PBC-2. The activation function is reLu for the hidden layers but in the output layer softmax activation function is used as it provides the results in terms of probability distribution for all outcomes. Back propagation with binary categorical cross entropy as loss function is adopted for the model.

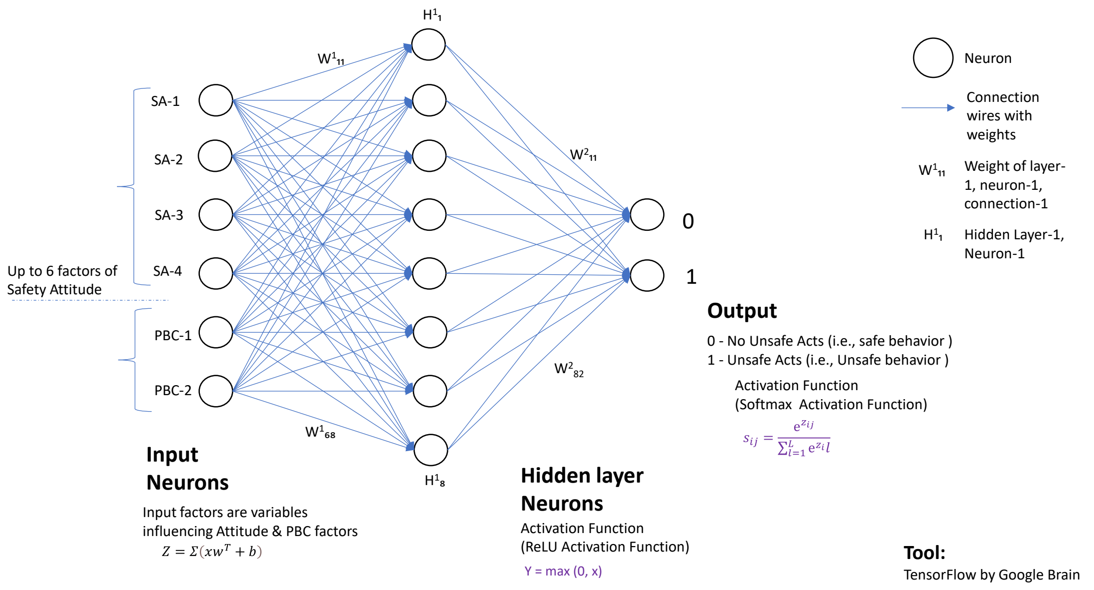

Figure 4.4: Initial architecture of ANN model

The optimizer used is Adam since it’s the most adapted and most efficient optimizer used in data science (Kingma and Ba, 2015). The tool used for deployment of model is Tensorflow with Keras API. Before adopting a high-level library for model training & validation, the ANN's backpropagation algorithm was scripted on python 3.7, without using any dedicated library source for better understanding of mathematical computataions of ANNs. The python pseudocode for learning basic computations is in appendix-1. Figure 4.4 shows the inputs, layers and outputs of initial structure.

4.5. Data collection

The instrument of data collection is a questionnaire. For the sample size, the rule-of-thumb is that it requires to be a 10 to 100 times the quantity of the features (Kavzoglu and Mather, 2003). With six (6) to ten (10) input features of SA and PBC, the acceptable sample size needs to be 100. One more rule-of-thumb is that the sample size requires to be a factor 50 to 1,000 times the quantity of prediction classes (Alwosheel, van Cranenburgh and Chorus, 2018). Since there are two prediction classes, sample size of at least 100 is required. A data set with 152 samples was taken. It took 20 weeks to collect data from eight different construction sites in Pakistan.

Following steps in Figure 4.5 are carried out for the collection of data.

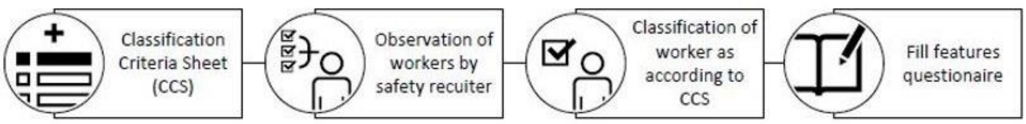

Figure 4.5: Data collection method

4.5.1. Classification Criteria Sheet (CCS)

It contained the criterias on the basis of which a worker is to be assigned a target class i.e., 0 being safe and 1 being unsafe. The CSS has the two criterias on the bais of which worker's should be classified. These are named as Performan criteria (P.C) 1 and 2. The PC-1 is the number of times a worker has violated a safety protocol whereas PC-

2 is the number of times took a shortcut to perform a task. These criterias are derived from unsafe act checklists from previous researches (Guo, Yiu and González, 2015; Goh *et al.*, 2018). The safety recruiters are provide with unsafe act checklists for their references Table 4-7. The CSS is attached in appendix-II.

Table 4-7: List of observed unsafe acts

Sr.	**Work Location**	**Unsafe Acts**
1	Working at height	Worker does not keep three points of contact with the ladder at all times.
		Worker carries any items while climbing up or down the ladder.
		Worker works on scaffolds, which are tagged unsafe, and/or not in safe condition.
		Safety harness is loosely worn.
		Scaffold component is removed or altered without approval from scaffold supervisor.
		Workers discards article/materials from height
		Worker works/rests near an opening, where there is a mentioned hazard of falling objects
		Taking unsafe shortcuts to access locations
		Distracted when performing task i.e. talking on mobile phone, eating, smoking
		Workers remove their helmets on worksite
2		Workers remove their helmets on worksite

	Foundation & Pits	Workers go up and down the foundation pit by establishing their own paths
		Workers cross dangerous area such as the edge of foundation pit and the sides of the gangway, etc.
		Distracted when performing task i.e. talking on mobile phone, eating, smoking
		Taking unsafe shortcuts to access locations
		Distracted when performing task i.e. talking on mobile phone, eating, smoking
		Workers stay too long at the edge of the foundation pit
3	Masonary works at ground level	Failure to ensure proper housekeeping
		Misuse of tools and equipment
		Stepping on guardrail to access unreachable site areas
		Distracted when performing task i.e. talking on mobile phone, eating, smoking
		Taking unsafe shortcuts to access locations

4.5.2. Observations by safety supervisor

Safety recuiters are the site supervisors of contractors which were requested to assist in the research work. Lusk et al. (1995) has described three methods to measure behaviour, which are: (1) observations, (2) supervisor's report, and (3) self-report. The self-reporting is the simplest technique but it has a level of biasness in it. Therefore, this study has used a mixed-approch where a safety supervisor is recruited to observe behaviors and report them. This technique has been used by reserachers for observation of unsafe acts in behaviour-based systems (Goh *et al.*, 2018; Guo, Goh and Le Xin Wong, 2018).

4.5.3. Classification of worker as per CCS

PC-1 and PC-2 were observed for 5 workers in a day. The maximum value of each criterion is assigned as 5; a worker ignored a safety protocol 5 times in a day. An overall score was computed by adding the scores of both criteria and a class is allotted to worker. The threshold for safe was kept at 4. That is if a worker had a total score 4 or less than 4, he was be classified as a safe worker (coded as class 0), otherwise he was put in unsafe class i.e., class 1.

4.5.4. Data on features questionnaire

After the observations, the worker's individual attributes were fetched on the questionnaire. The workers were informed about the confidentiality of their data through a consent form. Each worker wass requested to rate the scale item on a 5-point Likert scale. The questionnaire was translated in native language, for the purpose of better understanding of workers (Goh *et al.*, 2018). The feature questionnaire is attached in appendix-III. Table 4-8 shows the information of workers with various demographic details.

Table 4-8: Information of observed respondents

Item		**Frequency**	**Percentage**
Project type	Residential	6	66.6
	Commercial	2	33.3
Respondent's age	18-36	83	54.60
	39-60	69	45.39

Education level	Primary	86	56.57
	Secondary	55	36.18
	Diploma	11	7.23
Work experience	>5	38	25
	6-15	71	46.71
	more than 15	43	28.28
Work location	Heights	67	44.07
	Foundation & pits	34	22.36
	Ground	51	33.55
Trades	Formwork	39	25.65
	Steel & concreting	29	19.07
	Plastering	18	11.84
	Plumber	13	8.55
	Electrics	7	4.60
	Paints	15	9.86
	Masonry	31	20.39

4.6. Model Execution

The initial configurated model with 8 inputs is executed using the Tensorflow Keras API. Following steps in Figure 4.6 were taken to achieve a best parameter model.

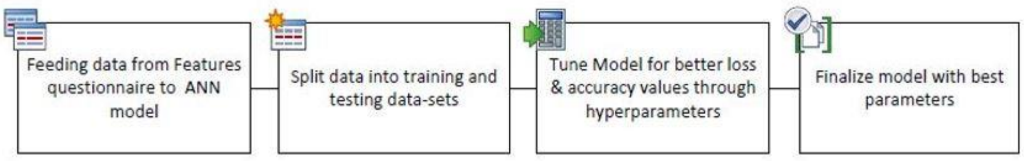

Figure 4.6: Model execution method

4.6.1. ANN input features

The data collected on feature questionnaire was feeded to the input layer of model. Since there were two measurement items for some of factors, therefore, a mean value was computed for each factor and then put as model input. (D. A. Patel & K. N. Jha, 2015). Figure 4.7 shows the working of Keras API for neural nets.

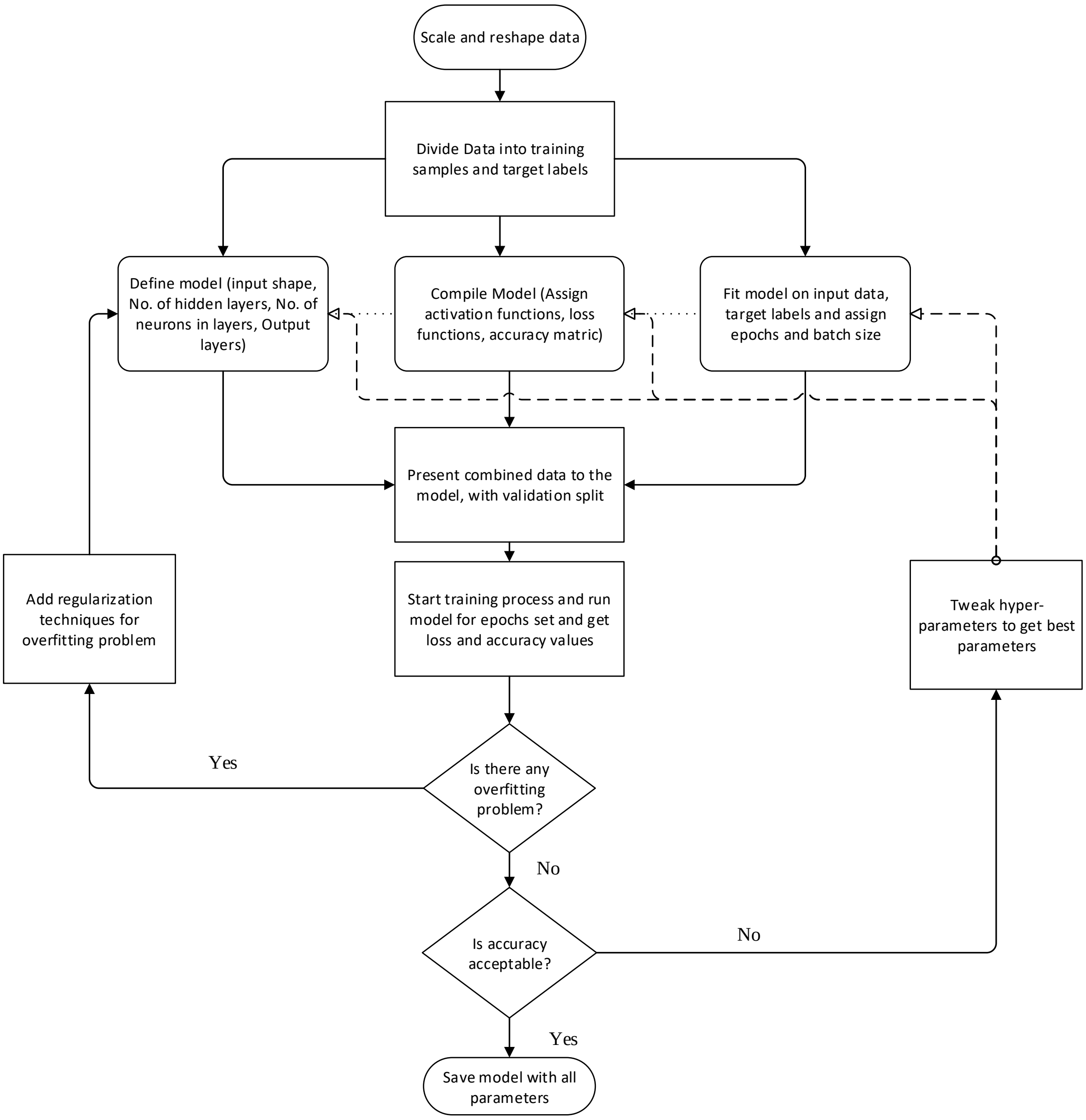

Figure 4.7: Schematic diagram of Keras API

4.6.2. Splitting datasets

Out of 152 data samples, 90% of data was used for training whereas 10% was kept aside for testing. Training dataset was for teaching the ANN model, the underlying patterns in dataset so that it can predict the outcome. In the Keras API, the training data set was further divided into two sets, one for training, which comprised 70 % of total training set, second was for validation, which comprised 30% of the training dataset. The validation dataset was to test the correctness of the model (Alpaydin, 2020).

4.6.3. Tuning

The initial model was run with multiple epochs and varying quantity of neurons in the hidden layer.

Table 4-9: Iterations & Tweaks on initial configurated model

Neurons in Hidden Layer	Batch Size	Epochs	Drop-out	TA	VA	TL	VL
11	32	100	0.2	0.69	0.72	0.42	0.47
12	32	100	0.2	0.88	0.87	0.28	0.4
13	32	100	0.2	0.94	0.83	0.2	0.53
14	32	100	0.2	0.92	0.7	0.181	0.5
15	32	100	0.2	0.94	0.88	0.18	0.43
16	32	200	0.2	0.87	0.77	0.19	0.49
17	32	200	0.2	0.87	0.77	0.26	0.44
18	32	200	0.2	0.9	0.77	0.17	0.56
19	32	200	0.2	0.92	0.77	0.2	0.51
20	32	200	0.2	0.89	0.83	0.22	0.49

Table 4-9 shows a few hit and trial runs with achieved loss and accuracy values. TA corresponds to training accuracy, VA is validation accuracy, TL is training loss and VL is validation loss values. After multiple runs, it was seen that the best training and validation accuracy this model achieves is around 0.8 and 0.75 repectively (Figure 4.8).

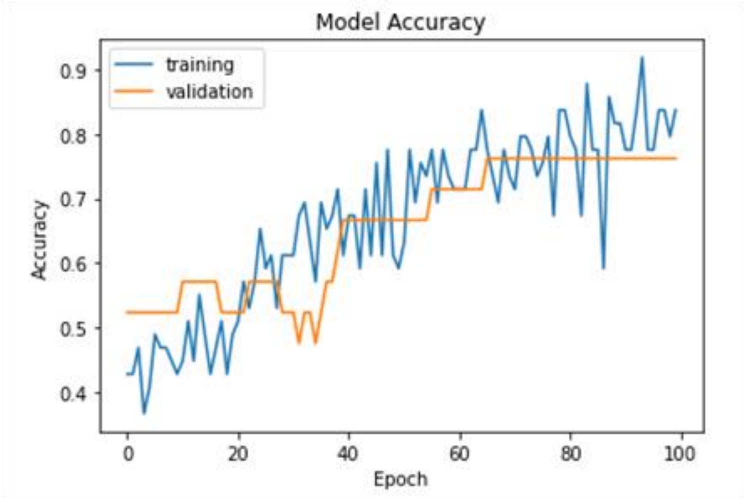

Figure 4.8: Training & validation accuracy of model with 1 hidden layer

The loss values were seen at 0.4 for training dataset and around 0.55 for validation data set Figure 4.9.

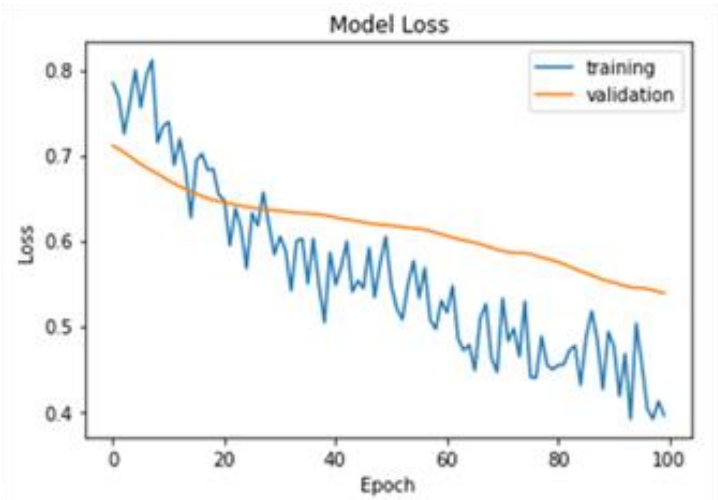

Figure 4.9: Training & validation loss for model with 1 hidden layer

Thus, the model was updated with more input neurons and a second hidden layer. Two (2) more factors of safety attitude were added in input layer and the second hidden layer was initially given four (4) neurons. A few iterations for this model are listed in Table 4-10.

Table 4-10: Iterations & Tweaks for model with 2 hidden layers

Epochs	Neurons in Hidden Layer 1	Neurons in Hidden Layer 2	Batch Size	Drop -out	TA	VA	TL	VL
50	14	4	32	0.3	0.87	0.84	0.31	0.5
100	15	5	32	0.3	0.82	0.86	0.12	0.42
200	16	6	32	0.3	0.93	0.83	0.18	0.34
300	17	7	32	0.3	0.97	0.67	0.042	0.85

From the above Table 4-10, it was concluded that model shows better performance with 6 neurons in second hidden layer. When epochs were 300 with seven neurons in second hidden layer, the model's validation loss increases to 0.85, making it an overfitted model.

4.6.4. Finalized model

Thus, during the process, the best-fitted model was achieved (Alwosheel et al., 2018) with a configuration of 10 input features, 1st hidden cover with 16 neurons and 2nd hidden layer with 6 neurons and 2 neurons in the output layer and are listed in Table 4-11 and ANN architecture shown in Figure 4.10. The loss and accuracy of the model are checked during the hyper tuning of model parameters.

Each time the model is run, it achieved training accuracy between 0.8 and 0.9; the validation accuracy reached is between 0.9 and 1 (Figure 4.11). Also, in every run, the loss values for training & validation are between 0.15 & 0.35 respectively (Figure 4.12).

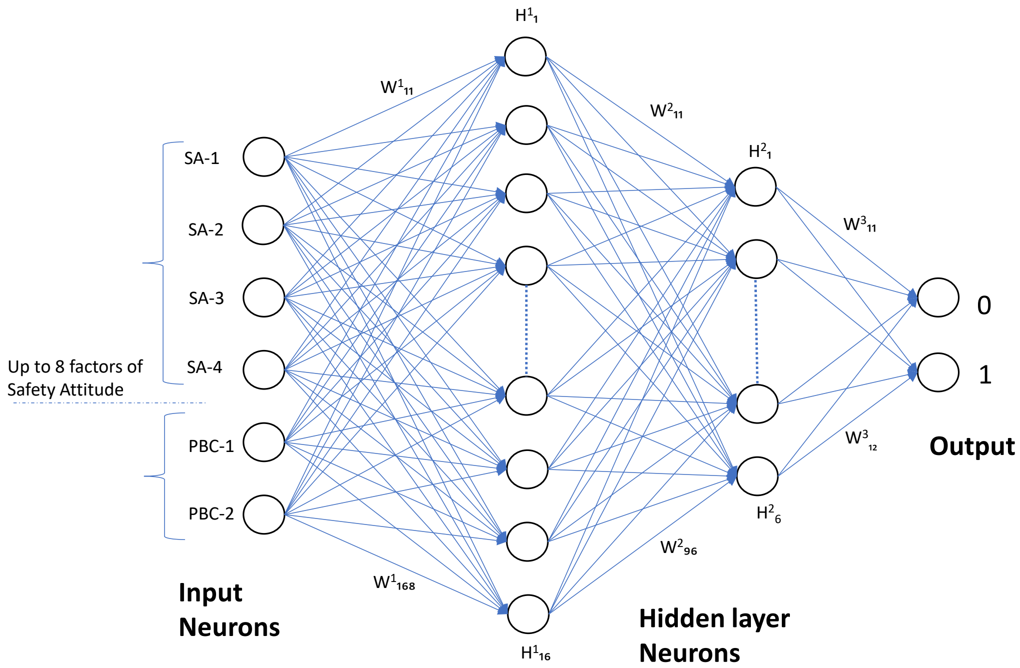

Figure 4.10: Final architecture of ANN model

Table 4-11: Parameters of finalized model

Epochs	Neurons in 1st hidden Layer	Neurons in 2nd Hidden Layer	Batch Size	Drop-out
200	16	6	32	0.2

The variation in the accuracy is due to the varying initial weights assigned every time the model is run from start. The model performed well on the validation dataset indicates that it generalized the problem well and is free from overfitting and underfitting problems.

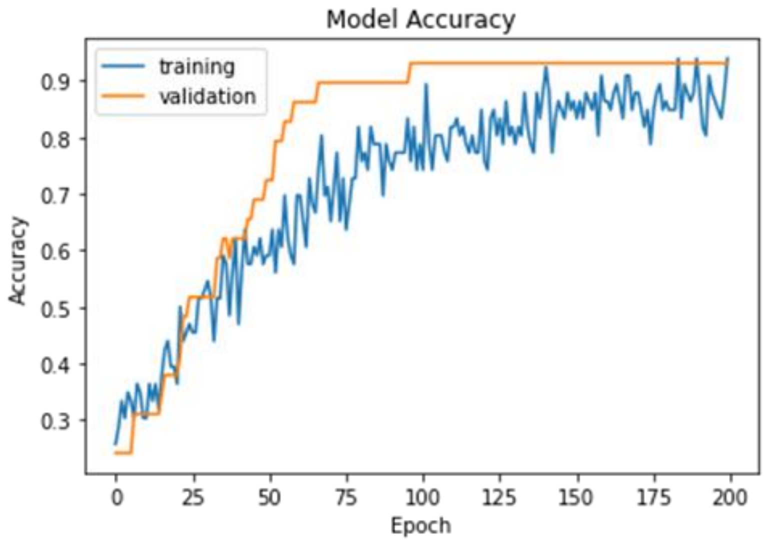

Figure 4.11: Training & validation accuracy of finalized model

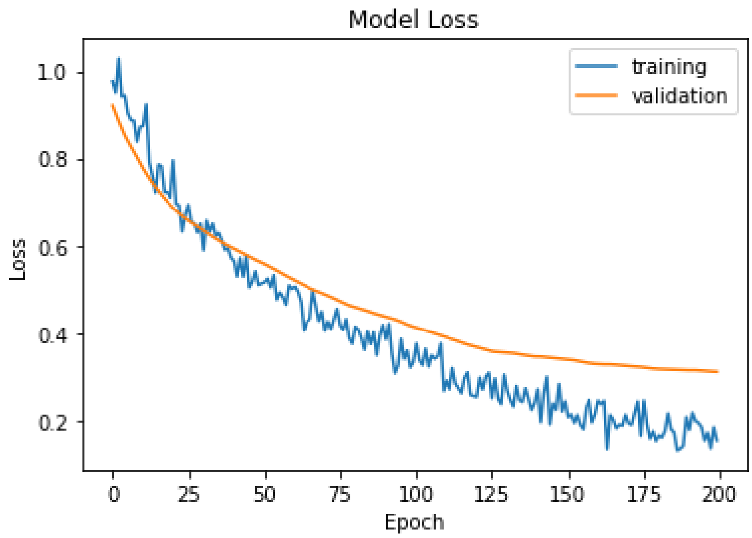

Figure 4.12: Training & validation loss of finalized model

4.6.5. Performance validation

The cross-validation technique, called as k-folds-cross validation is used that is particularly useful for smaller data sets. This is because when datasets are small, the number of samples in validation set is very low, and the resulting accuracy may not be credible. During this technique, instead of the holdout approach where data is split into fixed percentages for training and validation, the whole data set is divided into 10 subsets and in each subset 10% of data is set for validation (Yadav and Shukla, 2016). When the data is split into 10 subsets, its called 10-folds-cross technique. In the 10-fold cross technique, each subset is used for training and validation. The process is repeated 10 times and accuracy and loss values are determined (Yadav and Shukla, 2016). From Figure 4.13 the training and validation accuracies achieved range from 0.85 to 0.95, while Figure 4.14 shows the loss values for 10-folds-cross technique.

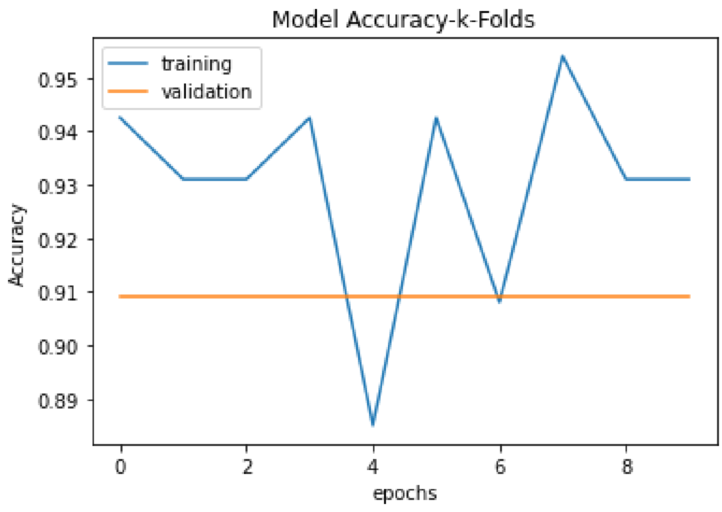

Figure 4.13: Training & validation accuracy of 10-cross-folds

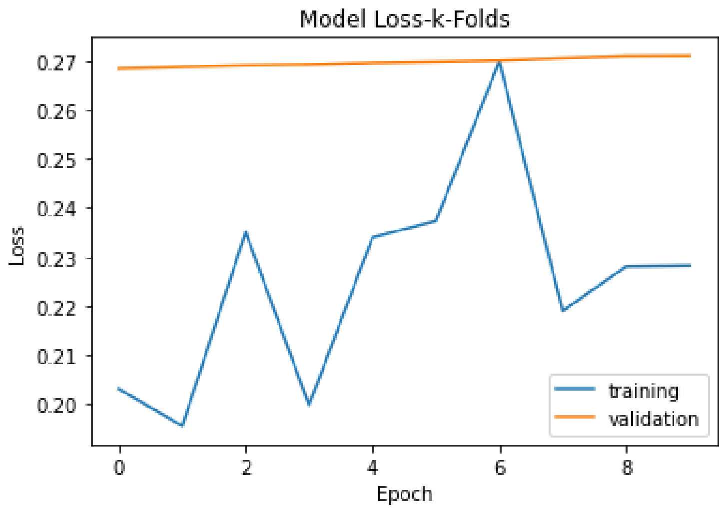

Figure 4.14: Training & validation loss of 10-cross-folds

4.6.6. Test dataset and accuracy

The 10% datatset set aside in the start was used to test the performance of model. 15 samples were in test dataset. Figure 4.15 shows the three steps taken for testing the model. At this point, target class is not provided to the model. The accuracy of model on test dataset is analyzed through a confusion matrix.

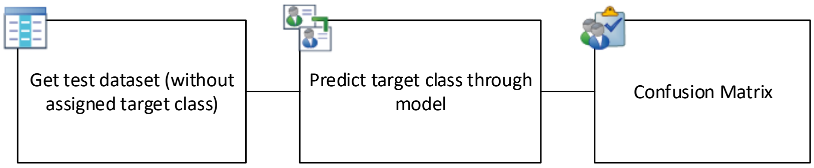

Figure 4.15: Method for testing the ANN model

4.6.6.1. Confusion matrix

A confusion matrix is used to check the functioning of the test dataset of the machine learning algorithm. The statistical parameters such as root mean square error (RMSE), mean absolute percentage error (MAPE), average absolute deviation (AAD), average percentage deviation (APD), and coefficient of variation (COV) are used to measure the performance of ANN regression algorithms (Patel and Jha, 2015b). This study used a classification approach; therefore, performance metrics is the confusion matrix. Visually, it is a table that outlines expected and test results and contrasts them with real-world values (Susmaga, 2004).

Table 4-12: Actual vs predicted class

SA-1	SA-2	SA-3	SA-4	SA-5	SA-6	SA-7	SA-8	PBC-1	PBC-2	Target	Predicted
3	3	4	3	3	2	2	3	2	3	1	1
3	5	4	4	4	3	3.5	3	3	3	0	0
3.5	2	2	2	2	2	3.5	3	2	3	1	1

3.5	3	3	3.5	2	3	3	2	3	2	1	1
3.5	3	3.5	3	3	3	2	3	3	3	1	1
3	5	4	4	4	3.5	3	4	3	3	0	0
3.5	3	2	2	3	2	3	3	3	3	1	1
4	4	3.5	4	3	3	3	3.5	3	3	1	0
3	3	3.5	2	2	3	3.5	3	3	2	1	1
3.5	3	4	3	4	3.5	4	4	4	3	0	0
3.5	3	3	3.5	2	2	3	2	3	3	1	1
3.5	3	3.5	4	3	3	4	3	4	3	0	0
3	5	4	2	2	3.5	3	3.5	3	3	1	1
3.5	3	2	2	3	2	3	3	3	3	1	1
4	4	4	4	4	3	3	4	4	3	0	0

Table 4-12 shows 15 samples of test dataset, with actual class and the predicted class by the model. There were 6 safe workers and 9 unsafe workers. The model was wrong in the prediction of one sample. Thus, the accuracy achieved was 93%. The confusion matrix in Figure 4.16 represents the same results in a confined tabular form.

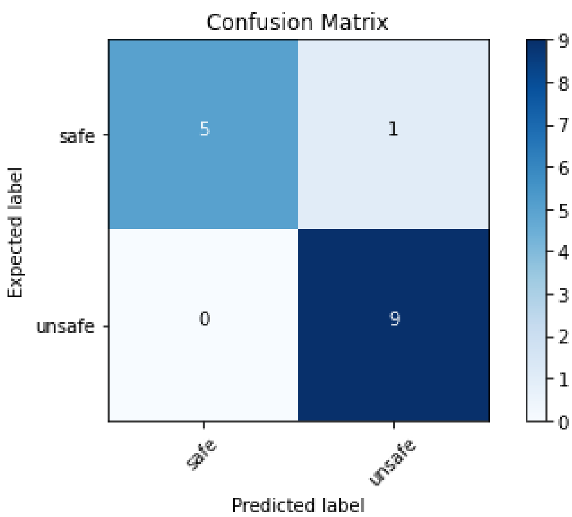

Figure 4.16: Accuracy visualization via confusion matrix

4.6.7. Prediction Engine

Prediction engines are the recommendation mechanisms that use machine learning algorithms to predict certain aspects using features, behaviors, and patterns of users (Kotu and Deshpande, 2019). Digital platforms are based on mass data collection and closed digital systems to predict content for users. The proposed propensity engine takes inputs from the users on a web server. The inputs are for features of behavioral constructs with measurement items i.e., SA & PBC, and must have values ranging from 1-5. Afterward, it predicts the propensity of the workers by classifying them into safe or unsafe workers. The propensity prediction engine was linked to the safety management program of the construction organizations.

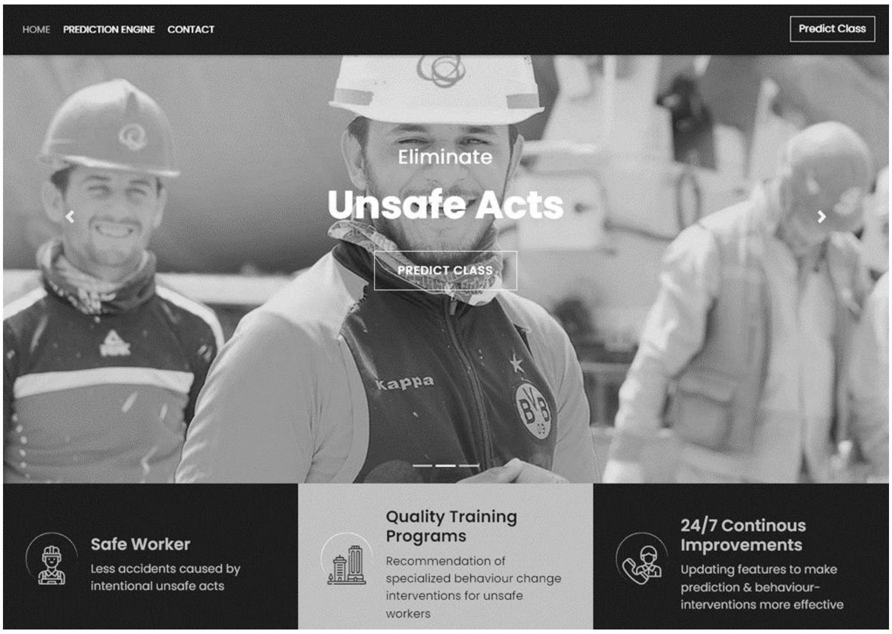

Figure 4.17: First page of proposed interface

Figure 4.17 shows the home page view of proposed user interface for the engine. Figure 4.18 shows the second tab of the engine, where individual factors are given and the user is asked to give input values for each factor on a 5-point likert scale. The interface is simple and contains the same measurement items used for data collection earlier. Figure 4.19 shows the predicted class by engine along with its general suggestion for the worker. In the example case, the engine predicted the worker to be of unsafe category and advises the safety managers to go for behavior interventions for the worker.

Figure 4.18: Second page with Individual factors on Likert scale

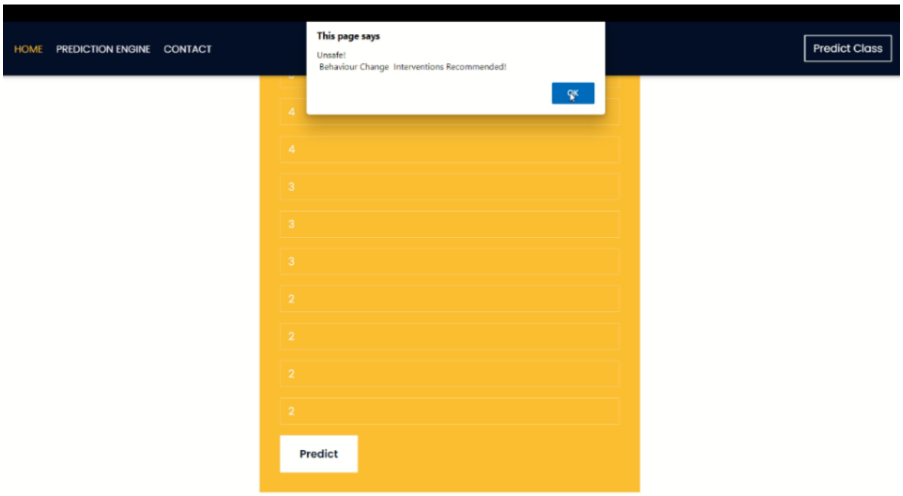

Figure 4.19: Suggestions by propensity engine

4.6.8. Framework for Adaptive-Behaviour Training

The study proposes an implementation framework where the propensity engine can be effectively used by the safety managers.

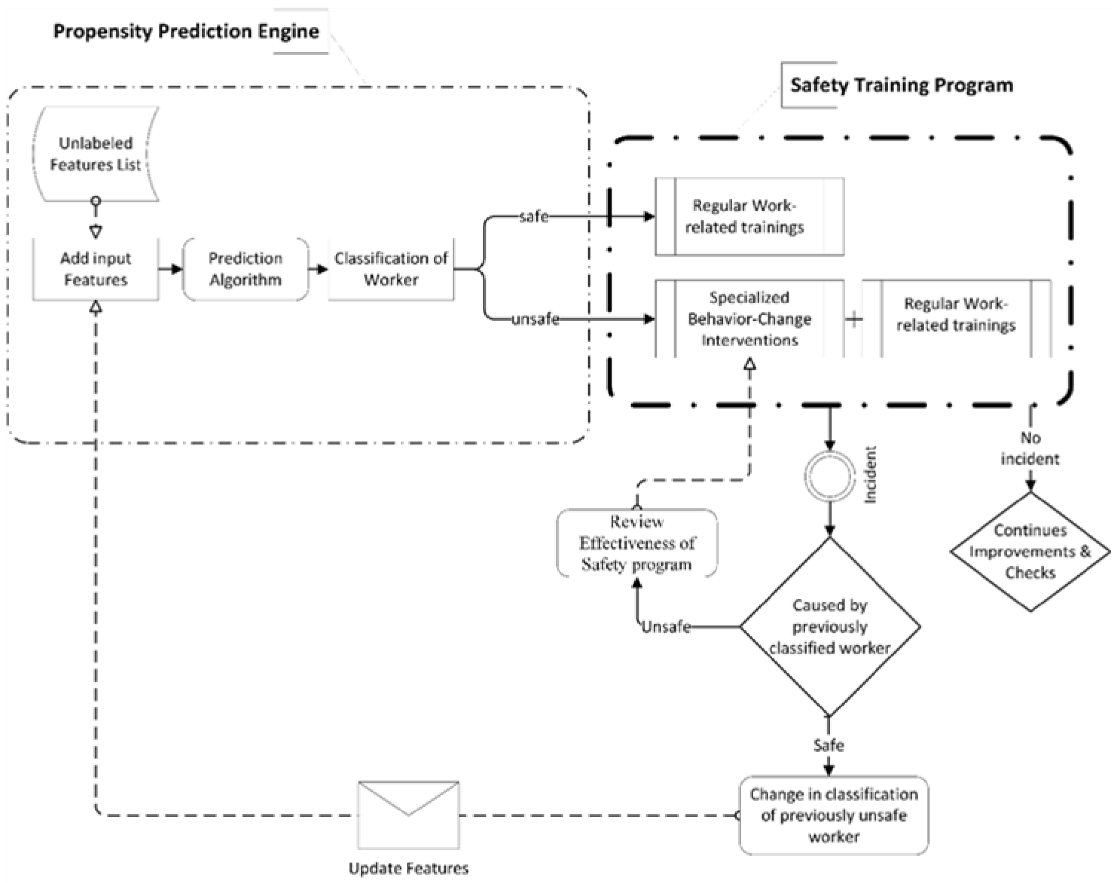

Figure 4.20: Propensity engine implementation framework

Figure 4.20 represents the framework for using the engine to improve safety training programs of construction organizations when the focus is to eliminate unsafe acts of workers through their behaviors. The framework suggests that once the workers are classified as safe or unsafe by the propensity engine, they should be given the appropriate training. The safe workers can be put for regular training regarding safe work protocols to avoid errors, while the unsafe workers are given behavior-change interventions along with the regular training. After training, if no incident is caused by unsafe behavior of workers, the management should keep on striving for continuous improvements in its safety program. On the other hand, if an incident occurs due to unsafe behavior, safety personnel should check if the involved person was previously classified as safe or unsafe. If worker was a safe one before, the features saved in the model's database should be updated and behavior-change interventions should be given again. If the incident was caused by an unsafe worker, the management can review the effectiveness of its behavior interventions and introduce new programs to cater to the problems in the previous one.

4.6.9. Modified HFACS-TPB for incident analysis:

When an incident takes place, if it is caused by the unsafe behaviour of worker, the procedure to identify causes of unsafe behaviour is explained through HFACS-TPB model. This modified accident investigation model has an additional item at level 4 i.e., Precursors of Unsafe Acts in Figure 4.21.

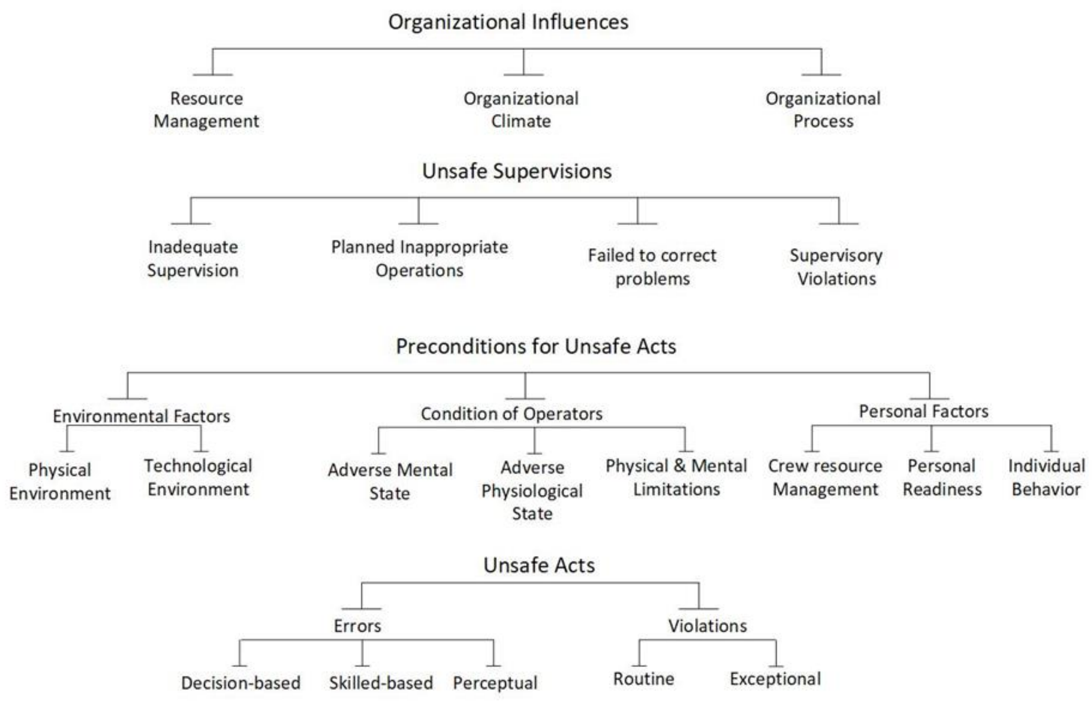

Figure 4.21: Proposed changes in HFACS model

The new item is named as Individual behavior. The individual behavior can be a a precursor to both errors and intentional acts. Since it is not possible to proactively predict errors, the model can be focused on exploring intentional acts only. The proposed model adds knowledge to accident causation model (HFACS) by proposing the addition of individual behavior is the precursor to intentional unsafe acts and the variations in individual behavior through can be explained through TPB.

This HFACS framework with behavior precursor provides an overall portraition of the primary causes of unsafe behaviours on construction sites. The proposed HFACS linked with TPB for predicting worker's proness to unsafe acts contributes to reseraches in Accident causation model and behaviour prediction models.

4.6.10. Chapter conclusion

This chapter included the explanation of each step of research methodology along with the findings. It included the steps to make the TPB-based theoretical foundation for the individual unsafe acts of workers, development of ANN model, training, validation and testing of the model. It further proposed an implementation framework for the integration of an unsafe act propensity prediction engine into organization's safety management systems.

Chapter 5

Conclusion

5.1. Discussion

The unsafe behavior of workers lead to unsafe acts and unsafe acts cause accidents. Although the causes of unsafe behaviors are varied but most of the time the direct cause is the personal preference of the individual. As those personal preference that are aligned with the safety work protocols do not put the environment at any risk but the personal preferences that negate the safety rules and regulations make the work environment vulnerable to accidents. From the SCM, it is suggested that the individual unsafe acts are the last defence against an accident and any fault in this barrier lead the hazard directly to mishap. It further says that if we put an extra safety barrier at organizational level, it is possible to stop the linear chain of unwanted events. The behavior modification based on the unsafe act propensity prediction model can act as a barrier at worker's level and also at top organizational level. From organizational point of view, if workers are assessed for their behavior is a defensive barrier against unsafe behaviors and it shows that organization is committed to eliminate such behaviors. Thus, if organization is determined to correct unsafe behaviors, it is analogous to putting a new safety barrier between organizational top level and the levels below them (in SCM). From the worker's level, the active levels nearest to a mishap in SCM, if the organization continousely thrive for better safety behavior of workers and is providing them with trainings and behavior change interventions, it makes the worker's strong enough to act as a barrier against any odds that come in their way. In simpler words, the behavior improvements make worker's resilient enough to resist any unsafe conditions that might have lead them into an unsafe act. Thus, the purpose of the unsafe act propensity prediction is to make the overall environment safe. The notion of assessing behavior and acts must not be taken as a way to blame workers for accidents. Its aim is to make the organization safety culture strong enough to not provide any loophole in its safety management systems. It should not be, in anycase, considered as a system for victim-blaming and punishing workers. This is only for improvements of behaviors and this improvement feature makes this research distinguished from BBS, where based on behavior, the reward or punishment is given to the workers.

5.2. Review of research objectives

The objectives of the study were:

1. Exlporation of the individual features that influence the behavior of construction workers and creation of a theoretical foundation
2. Development of a propensity prediction model for classification of construction worker's inclination towards
3. Theoretical frameworks for dedicated training programs integrated with propensity engine and modified HFACS for continues improvement in behaviour of workers.

5.3. Summary

Unchecked unsafe behaviours should be considered as unsafe conditions tolerated at organizational level and and there must be a system to assess behaviours of works for reducing unsafe acts at worker's level (Love and Smith, 2016). Assessing a worker's proneness to unsafe acts as a behaviour management strategy can help management in effectively managing intentional unsafe acts caused by unsafe behaviours. For predicting the inclination towards unsafe acts, this study adopted the Theory of planned behaviour (TPB) (Ajzen and Madden, 1986). The study has utilized two main constructs from TPB which can predict the behaviour of a person to make a predictive model using the Artificial Neural Networks (ANN). The individual factors which influence the behavioral construct were explored and put as inputs into ANN model. The model has successfully predicted the propensity for unsafe acts and it also recommends the behaviour interventions for unsafe workers. It is a multi-disciplinary study which has used Swiss cheese model (SCM) and Human Factor Analysis & Classification system (HFACS) to describe the problem of unsafe behavior of workers tolerated by organizations , Theory of planned behavior (TPB) for identification of features influencing behavior and provides a propensity prediction engine built on Artificial Neural Networks (ANN) to proactively address unsafe behaviors at organizational level. The study also provides investigation of incidents caused by unsafe acts via HFACS-TPB framework. This HFACS-TPB framework with behavior precursor provides an overall portraition of the primary causes of unsafe behaviours on construction sites. The proposed HFACS linked with TPB for predicting worker's proness to unsafe acts contributes to reseraches in accident causation models and behaviour prediction models.

5.4. Limitations

There are various limitations of the study. Firstly, the individual features under behavioral constructs do not take into account some complex factors such as personality traits. Also, the study doesnot take into account the geographical and economic factors which can also impact a worker's behaviour. Second limitation of the study is the data collection via safety recruiters as in this technique the chances of human biaseness are always present. The accuracy of safety recruiters while observing workers can also be seen as a limitation of study. Also, it can also be pointed out that behaviour can change for a short period of time as the presence of a an observer might alert workers to change their way of works. Third limitation is the less sample size. Although the criterias were fulfilled but larger datasets can yield models with better performances.

5.5. Future Studies

The future researches should can be administeres to cater these limitations. The reserachers can use cameras to monitor the activities of workers instead of relying on observations made by safety recruiters. Further, more features and record of previous behaviours of workers can be used to make a better prediction engine. Also, the true artificial intellidence (AI) with reinforcement learning can be implemented for behaviour prediction through continouse addition and learning of predictive model with datasets from all around the world, in a universal unsafe behaviour prediction database. These engines then can be integrated into Human Resource Management systems (HRMS) of construction organizations for behaviour-based hiring, training and behaviour-oriented company evaluations. The relationships between individual features with unsafe behaviours can be explored through statistical techniques such as confirmatory factor analysis and structural equation modelling.

Appendix-1: Pseudo-Code ANN Model

Terminologies:

n1 number of output neurons of layer_1 = 3

n2 number of output neurons of layer_2 = 3

Xd Dimension of input dataset = 2

X input data where different data sets and columns are values of parameters of the data set. In our case we have 300 input data sets with 2 variables in each.

W is weight matrix,

$$\begin{matrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{matrix}$$

$$\hat{y}_{i,j} = matrix\ with\ predicted\ value\ of\ y$$

Number of columns= Xd

Number of rows in W matrix = n1 or n2

w_{ij} gives weight for ith sample input and jth neuron

Z = matrix of summated outputs

s_{ij} = Predicted values (softmax output of ith sample)

L_i is loss of i^{th} sample

FP = Forward Pass

BP = Backward Pass

Forward pass steps (FP)

FP Step 1: Operations at Layer 1:

1. First, we initialize 1st layer with input variables (2) and number of neurons in first layer (3).
2. Once initialized the layer creates a random matrix of weights (n1 x Xd) and bias of size (n1xXd).

3. After, we apply linear equation and get Zij, which is the xw+b output for ith dataset and jth neuron.

$$Z = XW^T + B$$

4. Original form is wx+b but in order to use all X at the same time, we use the equation.

$$Z = \Sigma(xw^T + b)$$

5. Once we get the $Z = \Sigma(xw^T + b)$, we put this matrix through ReLu_Activation function. Equation for ReLu is

$$Y = \max(0, x)$$

FP Step 2: Operations at Layer 2:

1. The output from ReLu activation function is forwarded to layer_2 as inputs.
2. The no. of input neurons & output neurons in 2nd layer are defined using init. of layer_dense class. The input variables n1 = 3 and output neurons n2 = 3.
3. First, we again apply the linear model (xw+b) to layer_2 and get the Z matrix.

$$Z = \Sigma(xw^T + b)$$

4. Once, we get the Z matrix for layer_2, we apply the softmax activation function to get the final output. The formula for softmax activation is

$$s_{ij} = \frac{e^{z_{ij}}}{\sum_{l=1}^{L} e^{z_i l}}$$

$S_{i,j}$ denotes j^{th} Softmax's output of i^{th} sample.

5. We use the softmax activation to get the probabilities of each output class. Since ReLu is applied to each neuron of output layer & each neuron is independent of the outcome of the other neurons, therefore ReLu cannot give us a probability distribution across a layer.

   exp_values = np.exp(inputs - np.max(inputs, axis = 1, keepdims = True))

6. In order to use Softmax, we need to normalize values to avoid overflow problem. Overflow problem appears when we exponentiate a larger value and runtime warnings occur. Overflow problem is eliminated by subtracting maximum value from each row (each data set).

7. Now the largest value before exponentiation will be 0 and other values will be negative. Exponentiating these normalized values will give us a range of values between 0 & 1.
8. Then, we apply the softmax formula and get the predicted probabilities of each output class.

   S_{ij} = exp_values / np.sum(exp_values, axis = 1, keepdims = True

FP Step 3: Loss calculation

1. Now, we calculate loss by using Loss Categorical Cross entropy. The formula for which is $$L_i = -\sum_j y_{i,j} \log(\hat{y}_{i,j})$$

L_i is loss of i^{th} sample. y is actual value & y-hat is predicted value from Softmax.

2. Values of $S_{i,j}$ are clipped to 1e-7 so that the lowest value will be 1e-7. This will prevent 0 loss value as log of 0 is undefined.

   y_pred_clipped = np.clip(y_pred, 1e-7, 1 - 1e-7)

3. These y_pred_clipped are now indexed by the range of samples we have (300), then we grab the index of y_true values. This means that from each sample we grab the correct confidence value using the indexing of y_true values.

   correct_confidences = y_pred_clipped[range (samples), y_true]

4. In other words, since out of 3 predicted values, only 1 is correct. Therefore, the index of correct class will be used and other 2 class values will be 0, we can simplify the formula

$$L_i = -\log(\hat{y}_{i,k})$$

   K is the index of correct class probability.

5. From above simplified formula, the output we get is the negative log of likelihoods.

   negative_log_likelihoods = -np.log(correct_confidences)

   we calculate negative log as we are not calculating ypred-yact, rather take negative log of the calculated probabilities. Remember the complement of probabilities is the error.

6. Now, we calculate the overall loss which is the mean value of all sample losses GIVES ONE VALUE OF LOSS

   data_loss = np.mean(sample_losses)

Backward pass steps (BP)

Once we have calculated a loss value, we will optimize it by updating weights & biases using the chain rule.

$$\frac{dL}{\mathrm{d}w} = \frac{dL}{d\hat{y}} \times \frac{\mathrm{d}S}{\mathrm{d}z} \times \frac{\mathrm{d}z}{\mathrm{d}(mul)} \times \frac{\mathrm{d}(mul)}{\mathrm{d}w}$$

We will take the derivative of each function w.r.t to its inputs.

BP step 1: Calculating $\frac{dL}{d\hat{y}}$:

- Derivative of loss function i.e., Loss Categorical Cross entropy is

$$\frac{\partial L}{d\hat{y}} = \frac{-y_{i,j}}{\hat{y}_{i,j}}$$

Loss function.backward $(\hat{y}_{i,j}, y_{i,j})$

- The derivative of this loss function with respect to its inputs (predicted values at the i-th sample, since we are interested in a gradient with respect to the predicted values) equals the negative ground-truth vector, divided by the vector of the predicted values (which is also the output vector of the softmax function).

self.dinputs = -y_true / dvalues

*(dividing each value in y_true vector by each value in y_pred vector)

- We're turning numerical labels into one-hot encoded vectors since we need vectors here, we'll use the np.eye method which, given a number, n, returns an n x n array filled with ones on the diagonal and zeros everywhere else.

y_true = np.eye(labels)[y_true]

- The second operation is the gradient normalization. Optimizers sum all of the gradient related to each weight and bias before multiplying them by the learning rate. What this means, in our case, is that the more samples we have in a dataset, the more gradient sets we'll receive at this step, and the bigger this sum will become. As a consequence, we'll have to adjust the learning rate according to each set of samples. To solve this problem, we can divide all of the gradients by the number of samples.

  self.dinputs = self.dinputs / samples

BP step 2: Calculating $\frac{dS}{d\hat{z}}$:

1. Derivative of loss function i.e., softmax is

   $$S_{i,j}\, \partial_{j,k} - S_{i,j}\, S_{i,k}$$

   The left part of this equation contains the Kronecker delta multiplied by softmax output.

2. The left part can be achieved by using np.diagflat method which creates an array using an input vector as the diagonal:

   np.diagflat(softmax_output)

3. The other part of the equation is $S_{i,j}\, S_{i,k}$ — the multiplication of the Softmax outputs, iterating over the j and k indices respectively. Since, for each sample (the i index), we'll have to multiply the values from the Softmax function's output (in all of the combinations), we use the dot product operation. For this, we'll just have to transpose the second argument to get its row vector form.

   np.dot(softmax_output, softmax_output.T)

   combining steps 2 & 3:

jacobian_matrix = np.diagflat(softmax_output) -np.dot(softmax_output, softmax_output.T))

4. The matrix result of the equation is called the Jacobian matrix. In our case, the Jacobian matrix is an array of partial derivatives in all of the combinations of both input vectors. We are calculating the partial derivatives of every output of the Softmax function with respect to each input separately.
5. The code implemented is:

   i. **Create uninitialized array,**

   self.dinputs = np.empty_like(dvalues)

   First, we created an empty array (which will become the resulting gradient array) with the same shape as the gradients that we're receiving to apply the chain rule. The np.empty_like method creates an empty and uninitialized array. Uninitialized means that we can expect it to contain meaningless values, but we can set them afterwards.

   ii. **We use a for loop and attain a single row from S_{ij} and multiply with itself followed by multiplication of S_{ik}**

   for index, (single_output, single_dvalues) in enumerate (zip (self.output, dvalues)):

   single_output = single_output.reshape(-1, 1)

   jacobian_matrix = np.diagflat(single_output) - np.dot(single_output, single_output.T)

The for loop, iterates the sample-wise pairs of the outputs and gradients, calculating the partial derivatives

It:

- takes single output of 1 sample (activation2.output) , resize it to shape (3x1).

```
[[0.33333333]
 [0.33333333]
 [0.33333333]]
```

- Takes single values from dvalues(loss_function.dinputs) , size 1x3,

```
[-0.01 -0.   -0.  ]
```

- Applies the dot product and subtract from Kronecker delta array & gives jacobian matrix.

```
[[ 0.22222222 -0.11111111 -0.11111111]
 [-0.11111111  0.22222222 -0.11111111]
 [-0.11111111 -0.11111111  0.22222222]]
```

- Now, we calculate the final product (applying the chain rule) of the Jacobian matrix and gradient vector (from the passed-in gradient array), storing the resulting vector as a row in the dinput array. We're going to store each vector in each row while iterating, forming the output array.

self.dinputs[index] = np.dot(jacobian_matrix, single_dvalues)

$$=\frac{dL}{d\hat{y}} x \frac{dS}{d\hat{z}}$$

```
[-0.00222222  0.00111111  0.00111111]
```

The for loop creates a single row of $\frac{dL}{d\hat{y}}$x$\frac{dS}{d\hat{z}}$ **matrix of size 1X3. This creates an array of 300 x 3 size.**

```
[[-0.00222222  0.00111111  0.00111111]
 [-0.00222226  0.00111118  0.00111109]
 [-0.0022223   0.00111124  0.00111106]
 [-0.00222234  0.00111131  0.00111103]
 [-0.00222237  0.00111135  0.00111103]
 [-0.00222231  0.0011113   0.00111102]
 [-0.00222244  0.00111148  0.00111095]
 [-0.00222227  0.00111123  0.00111104]
 [-0.00222254  0.00111164  0.0011109 ]
 [-0.00222252  0.00111155  0.00111097]
 [-0.00222214  0.0011112   0.00111095]
 [-0.00222264  0.00111177  0.00111087]
 [-0.00222254  0.00111156  0.00111098]
 [-0.00222264  0.00111174  0.00111091]
 [-0.00222208  0.00111127  0.00111081]
 [-0.00222265  0.00111172  0.00111093]
 [-0.00222247  0.00111153  0.00111095]
 [-0.0022226   0.00111162  0.00111097]
 [-0.00222266  0.00111171  0.00111095]
```

BP step 2: Calculating $\frac{\mathbf{dz}}{\mathbf{d}(\boldsymbol{mul})} \times \frac{\mathbf{d}(\boldsymbol{mul})}{\mathbf{d}\boldsymbol{w}}$:

i. Since the partial derivative of summated function will always be 1, so value of $\frac{\mathbf{dz}}{\mathbf{d}(\boldsymbol{mul})} = \mathbf{1.}$

ii. The second part is the partial derivative of multiplication function w.r.t weights which equals inputs. Therefore, we multiply the derivative values from previous values to get the gradient matrix of weights. Size (3,3)

self.dweights = np.dot(self.inputs.T, dvalues)

(3,300) x (300,3) , since we defined 3 inputs in layer_2

```
[[ 0.00093682 -0.0011407   0.00020388]
 [-0.000561   -0.00059865  0.00115965]
 [ 0.00045724 -0.00112551  0.00066827]]
```

iii. Similarly, we get the dinputs array of shape (300,3) self.dinputs = np.dot(dvalues, self.weights.T)

```
[[ 3.86585688e-05  1.87701860e-04  4.10558407e-05]
 [ 3.86600768e-05  1.87703145e-04  4.10567395e-05]
 [ 3.86804631e-05  1.87693945e-04  4.10657966e-05]
 [ 3.87131953e-05  1.87682114e-04  4.10806810e-05]
 [ 3.86855908e-05  1.87692794e-04  4.10682102e-05]
 [ 3.86844469e-05  1.87693678e-04  4.10677447e-05]
 [ 3.87632728e-05  1.87664023e-04  4.11034537e-05]
 [ 3.87784864e-05  1.87658528e-04  4.11103723e-05]
 [ 3.87938519e-05  1.87652979e-04  4.11173601e-05]
 [ 3.88156138e-05  1.87645122e-04  4.11272569e-05]
 [ 3.87869746e-05  1.87655463e-04  4.11142325e-05]
 [ 3.88598878e-05  1.87629142e-04  4.11473922e-05]
 [ 3.88762416e-05  1.87623241e-04  4.11548299e-05]
```

iv. Partial derivatives of biases is obtained by summation of all dvalues across the column, and resultant vector of size (1x3) is obtained.

self.dbiases = np.sum(dvalues, axis = 0 , keepdims = True)

```
[[-0.00057449  0.00089898 -0.00032449]]
```

Getting partial derivatives values of weights of layer_1:

i. The dinput values from layer_2.backward are propagated to activation_1.backward i.e., ReLu function.

activation1.backward(layer_2.dinputs)

ii. From dinputs, where the input value is negative, the value will become 0.

self.dinputs[self.inputs <= 0] = 0

```
[[ 0.00000000e+00  0.00000000e+00  0.00000000e+00]
 [ 0.00000000e+00  1.87703145e-04  4.10567395e-05]
 [ 3.86804631e-05  0.00000000e+00  0.00000000e+00]
 [ 3.87131953e-05  0.00000000e+00  0.00000000e+00]
 [ 3.86855908e-05  0.00000000e+00  4.10682102e-05]
 [ 3.86844469e-05  0.00000000e+00  4.10677447e-05]
 [ 3.87632728e-05  0.00000000e+00  0.00000000e+00]
 [ 3.87784864e-05  0.00000000e+00  0.00000000e+00]
 [ 3.87938519e-05  0.00000000e+00  0.00000000e+00]
 [ 3.88156138e-05  0.00000000e+00  0.00000000e+00]
 [ 3.87869746e-05  0.00000000e+00  0.00000000e+00]
 [ 3.88598878e-05  0.00000000e+00  0.00000000e+00]
```

iii. From this, we get a new array of (300,3). Which is further propagated backwards into layer_1.

layer_1.backward(activation1.dinputs)

iv. From the backward method in layer_dense class, the gradients of weights are obtained (as done previously to get the dweights & dbiases of layer_2).

self.dweights = np.dot(self.inputs.T, dvalues)

(2,300) x (300,3) , since we defined 3 inputs in layer_1

print(layer_1.dweights)

```
[[-0.00189918 -0.00072037  0.00191323]
 [-0.00286301  0.00176763  0.00050791]]
```

self.dbiases = np.sum(dvalues, axis = 0 , keepdims = True)

```
[[-0.00170699 -0.00378996 -0.00106153]]
```

Optimization: Adam

Once we have calculated the gradient, we can use this information to adjust weights and biases to decrease the measure of loss. We will do this with the help of Adam optimizer. All optimizers are just variants of **stochastic gradient descent (SGD).**

ADAM:

Adam , short for **Adaptive Momentum** , is currently the most widely-used optimizer and is built atop RMSProp, with the momentum concept from SGD added back in. It uses the concepts of momentum from SGD and Cache from RMSProp (Root Mean Squared propagation). In order to understand it, we have to understand the concept of momentum & Cache.

The equation for Adam is:

$$\Delta w_i(t) = -\frac{\eta}{\sqrt{G_I(t) + \in}} M_i(t)$$

Where,

Mi (t) is momentum from SGD

Gi (t) is adaptive function(cache) from RMSProp.

The $\in$ is a hyperparameter (pre-training control knob setting) preventing division by 0.

The epsilon value is usually a small value, such as 1e-7 , which we'll be defaulting to.

- **Momentum:** Momentum creates a rolling average of gradients over some number of updates and uses this average with the gradient at each step.

$$M_i(t) = \alpha M_i(t-1) + (1-\alpha)\frac{\partial L}{\partial w_i}(t)$$

Easy understanding of momentum:

Another way of understanding this is to imagine a ball going down a hill even if it finds a small hole or hill, momentum will let it go straight through it towards a lower minimum the bottom of this hill. This can help in cases where we are stuck in some local minimum (a hole), bouncing back and forth. With momentum, a model is more likely to pass through local minimums, further decreasing loss. Simply put, momentum may still point towards the global gradient descent direction.

- **Cache**: RMSProp provides a way to normalize parameter updates by keeping a history of previous updates. This history is in Cache. Its also called Adaptive Magic function. Each individual weight in the whole network keeps track of its own G function to normalize its own steps. The formula is :

$$G_I(t) = \beta\, G_I(t-1) + (1-\beta)(\frac{\partial L}{\partial w_i}(t))^2$$

cache = rho * cache + (1 - rho) * gradient ** 2

here β and (1- β), (rho in code) are assigned weights, (also called as Beta_2 in next paragraph) which effectively keep track of moving averages of G-values. Beta is a hyperparameter, with value ranging from 0 &1.

The Adam optimizer additionally adds a bias correction mechanism. (Do not confuse this with the layer's bias). The bias correction mechanism is applied to the cache and momentum, compensating for the initial zeroed values before they warm up with initial steps. To achieve this correction, both momentum and caches are divided by *1-coefficient **step.*

***Corrected Momentum**:*

$$\frac{M_i}{1 - \alpha^{step}}$$

***Corrected Cache**:*

$$\frac{G_i}{1 - \beta^{step}}$$

Step for coding of Adam:

i. Create momentum arrays & Cache arrays using np.zeros_like(). (weights & biases)

layer.weight_momentums = np.zeros_like(layer.weights)

layer.weight_cache = np.zeros_like(layer.weights)

layer.bias_momentums = np.zeros_like(layer.biases)

layer.bias_cache = np.zeros_like(layer.biases)

ii. Now update momentum with current gradients the formula,

$$M_i(t) = \alpha M_i(t-1) + (1-\alpha)\frac{\partial L}{\partial w_i}(t)$$

#α in code is written as beta_1 , which is initialized at a value of 0.9

```
layer.weight_momentums =
self.beta_1 * layer.weight_momentums + ( 1 - self.beta_1) * layer.dweights
layer.bias_momentums =
self.beta_1 * layer.bias_momentums + ( 1 - self.beta_1) * layer.dbiases
```

iii. Now, we apply the correction mechanism to get corrected momentums of weights & biases,

$$\frac{M_i}{1 - \alpha^{step}}$$

#Self.iterations starts at 0, but we need to start with 1.

```
weight_momentums_corrected =
    layer.weight_momentums / ( 1 - self.beta_1 **(self.iterations + 1 ))

bias_momentums_corrected =
    layer.bias_momentums / ( 1 - self.beta_1 ** (self.iterations + 1 ))
```

iv. Now we update cach with squared current gradients,

$$G_I(t) = \beta\, G_I(t-1) + (1-\beta)(\frac{\partial L}{\partial w_i}(t))^2$$

#β in code is written as beta_2 , which is initialized at a value of 0.999

```
layer.weight_cache =
self.beta_2 * layer.weight_cache + ( 1 - self.beta_2) * layer.dweights ** 2
```

layer.bias_cache =

self.beta_2 * layer.bias_cache + (1 - self.beta_2) * layer.dbiases ** 2

v. Now, we apply the correction mechanism to get corrected cache of weights & biases,

$$\frac{G_i}{1-\ \beta^{step}}$$

weight_cache_corrected =

layer.weight_cache / (1 - self.beta_2 ** (self.iterations + 1))

bias_cache_corrected =

layer.bias_cache / (1 - self.beta_2 ** (self.iterations + 1))

vi. Finally we update the layer weights & biases using Adam's formula,

$$\Delta w_i(t) = \ -\frac{\eta}{\sqrt{G_I(t)}+\in}\ M_i(t)$$

layer.weights += - self.current_learning_rate * weight_momentums_corrected / (np.sqrt(weight_cache_corrected) + self.epsilon)

layer.biases += - self.current_learning_rate * bias_momentums_corrected / (np.sqrt(bias_cache_corrected) + self.epsilon)

Appendix-II: Classification Criteria Sheet

Guidelines:

1. **Worker id**: Worker's name and trade to be combined to form a unique code, which will keep worker's identity as unnamed. It is done for the purpose of keeping data of worker as confidential. So that he can feel free to answer the questions of features questionnaire.
2. **Performance Criteria:** The workers are to be classified on the basis of two criteria's: P.C-1 and P.C-2.
   i. **P. Criteria-1**: No of times a worker has ignored a safety protocol. e.g., removed PPE.
   ii. **P. Criteria-2**: No of times a worker took an unsafe shortcut e.g., jumping from formworks to get to the ground.
3. **Total Score**: P.C-1 and P.C-2 observed in a day by safety supervisor. The maximum value a of each criterion is assigned is 5. E.g., a worker has ignored a safety protocol 5 times in one day.
4. **Total score:** It is the sum of the values from both criteria.
5. **Allotted Class**: 0 represents a safe worker and 1 represents an unsafe 1. Class is allotted on the basis of total score. The threshold for safe is 4. That is if a worker has a total score 4 or less than 4, he will be classified as a safe worker (coded as class 0), otherwise he will be put in unsafe class i.e., class 1.

Worker ID:	P. Criteria-1	P. Criteria-2	Total score	Allotted Class
1-FW	3	3	6	1

In the above example:

i. **1-FW** = Worker 1 working on Formworks.

ii. **P. Criteria-1 = 3**, worker ignored a safety protocol 3 times in a day.

iii. **P. Criteria-2 = 3**, worker did work in an unsafe manner 3 times a day.

iv. **Total score = 6,** sum of both criteria.

v. **Allotted Class = 1,** sum is greater than 4, worker is classified as unsafe.

Note:

- **Maximum 5 workers are to be observed in a day.**
- **Mention date on each table.**

Date:

Worker ID:	**P. Criteria-1**	**P. Criteria-2**	**Total score**	**Allotted Class**

Date:

Worker ID:	**P. Criteria-1**	**P. Criteria-2**	**Total score**	**Allotted Class**

Date:

Worker ID:	P. Criteria-1	P. Criteria-2	Total score	Allotted Class

Appendix-III: Survey Questionnaire

Consent to Participate in a Research Project

Survey Project: ---

Administering Organization: National University of Sciences & Technology, Islamabad.

Investigator: Rafia Nawaz (MSc Research Student)

General Information:

You are invited to participate in a research study being conducted at the Department of Construction Engineering & Management, NUST, Islamabad. The purpose of this study is to explore personal factors which influence behavior. Your participation in this study is entirely voluntary and you may withdraw at any time.

Consent to participate:

- I agree to participate in the above research as described in the information statement.
- I can withdraw from the project at any time and do not have to provide any reason for doing so.
- I understand I only have to answer questions that I want to.
- I understand that my name and my identity information will be withheld to protect participant anonymity.

- I have the opportunity to ask any questions regarding the research from the researcher and these questions will be answered to my satisfaction.

Signing this document means that the above information has been described to you orally and that you voluntarily agree to participate.

Signature of the participant: ___________________ Date: ____________________

Study Contact:

Rafia Nawaz

Research Student, MSc Construction Engineering & Management,

School of Civil & Environmental Engineering, NUST-12, Islamabad.

Email: rnawaz.cem19@student.nust.edu.pk

General Assumptions peculiar to this research

The study has the following assumptions:

1. Respondent has received primary-level education.
2. Respondent has been given necessary safety training.
3. Respondent's age is between 18-50 yrs.
4. Respondent is provided with safety equipment necessary for his type of task.
5. Respondent has work experience of min. 2 years.

Questionnaire

Part I; Respondent's General Information:

Please enter your name, position, and the details of your organization.

1. Date of survey:
2. Name of respondent:
3. Age:
4. Education level:
5. Experience (years):
6. Work Location:
7. Position/Trade:
8. Organization:
9. Project Type:

Part II; Respondent's overall behavior as reported by Site supervisor/ Safety supervisor: (Check box)

Safe	**Unsafe**

Part III; Underlying factors influencing overall behavior

Factors influencing the attitude of workers:

Please indicate your level of disagreement/agreement with the following statements about construction site safety with the use of five-point Likert scales.

Q#1: PR-a: I think the likelihood of falling from height and getting injured is more.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#2: PR-b: I think the possible consequences of fall accidents are severe.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#3: WE: From my own work experience on construction sites, I think safety procedures implementation is important.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#4: SK-a: I am aware of risks associated with my job and health safety.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#5: SK-b: I am aware of necessary precautions to be taken while doing the job at construction sites.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#6: WEU-a: I think Personal protective equipment (PPEs) is effective for our safety at construction sites.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#7: WEU-b: I think avoiding hazards in our work environment is possible.

1	2	3	4	5

Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#8: HA: I can identify potentially hazardous situations.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#9: A-a: I believe that anyone can be involved in an accident either young or elder.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#10: A-b: I believe that everyone on the worksite either young or elder should comply with safety rules.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#11: B-a: I believe that an accident should not be regarded as an act of nature as it is possible to avoid accidents by complying with safety rules.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#12: B-b: I believe that my safety is my own responsibility.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#13: PB-a: I believe that most of the safety procedures are convenient.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Q#14: PB-b: I believe that all rules and policies relevant to my job are practical.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

Factors of perceived behavioral control:

Please indicate your level of agreement with the following statements about construction site safety with the use of a 5-point easy/difficult scale given along each statement.

Q#15:CM-a: For me to work at height without any PPE would be... (very easy/very difficult).

1	2	3	4	5
Very easy	Easy	Neutral	Difficult	Very difficult

Q#16: CM-b: I believe that I can…… perform my work without having an accident. (definitely/definitely do not)**.**

1	2	3	4	5
Definitely	somehow	Neutral	Can not	Definitely not

Q#17: AM-a: I have …. over my task without a PPE. (Complete control- No control).

1	2	3	4	5
Complete control	More control	Neutral	Less control	No control

Q#18: AM-b: I believe that the number of external events outside my control which can cause an accident while doing work at height is.... (None-Numerous)

1	2	3	4	5
None	Very few	Neutral	few	Numerous

REFERENCES

Abd, A. M. and Naseef, F. S. (2019) 'Predicting the Final Cost of Iraqi Construction Project Using Artificial Neural Network (ANN)', *Indian Journal of Science and Technology*, 12(28), pp. 1–7. doi: 10.17485/ijst/2019/v12i28/145640.

Abubakar, A. M. *et al.* (2019) 'Applying artificial intelligence technique to predict knowledge hiding behavior', *International Journal of Information Management*, 49(October 2018), pp. 45–57. doi: 10.1016/j.ijinfomgt.2019.02.006.

Ajzen, I. (2002) 'Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior', *Journal of Applied Social Psychology*, 32(4), pp. 665–683. doi: 10.1111/j.1559-1816.2002.tb00236.x.

Ajzen, I. and Fishbein, M. (2010) 'Predicting Changing Behavior', *Taylor & Francis Group*, p. 518.

Ajzen, I. and Madden, T. J. (1986) 'Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control', *Journal of Experimental Social Psychology*, 22(5), pp. 453–474. doi: 10.1016/0022-1031(86)90045-4.

Akhavian, R. and Behzadan, A. H. (2016) 'Productivity Analysis of Construction Worker Activities Using Smartphone Sensors', *16th International Conference on Computing in Civil and Building Engineering (ICCCBE)*, pp. 1067–1074.

Aksorn, T. and Hadikusumo, B. H. W. (2007) 'The unsafe acts and the decision-to-err factors of Thai construction workers', *Journal of Construction in Developing Countries*, 12(1), pp. 1–25.

Al-Sobiei, O. S., Arditi, D. and Polat, G. (2005) 'Predicting the risk of contractor default in Saudi Arabia utilizing artificial neural network (ANN) and genetic algorithm (GA) techniques', *Construction Management and Economics*, 23(4), pp. 423–430. doi: 10.1080/01446190500041578.

Almeida, A. and Azkune, G. (2018) 'Predicting human behaviour with recurrent neural networks', *Applied Sciences (Switzerland)*, 8(2). doi: 10.3390/app8020305.

Alpaydin (2020) *Introduction to Machine Learning, (Adaptive Computation and Machine Learning series)*. Fourth. The MIT Press.

Alpaydin, E. (2020) *Introduction to Machine Learning (Adaptive Computation and Machine Learning series)*. fourth edi. The MIT Press.

Alwosheel, A., van Cranenburgh, S. and Chorus, C. G. (2018) 'Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis', *Journal of Choice Modelling*, 28(July), pp. 167–182. doi: 10.1016/j.jocm.2018.07.002.

Aniekwu, N. (2007) 'Accidents and safety violations in the Nigerian construction industry', *Journal of Science and Technology (Ghana)*. doi: 10.4314/just.v27i1.33027.

Arciszewski et al. (1995) 'Construction Induction: The key to design creativity'.

Attalla, M. and Hegazy, T. (2003) 'Predicting Cost Deviation in Reconstruction Projects: Artificial Neural Networks versus Regression', *Journal of Construction Engineering and Management*, 129(4), pp. 405–411. doi: 10.1061/(asce)0733-9364(2003)129:4(405).

Baldissone, G. *et al.* (2019) 'The analysis and management of unsafe acts and unsafe conditions. Data collection and analysis', *Safety Science*, 119(September 2017), pp. 240–251. doi: 10.1016/j.ssci.2018.10.006.

Bangaru, S. S. *et al.* (2021) 'ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors', *Automation in Construction*, 126(March), p. 103653. doi: 10.1016/j.autcon.2021.103653.

Baysari, M. T., McIntosh, A. S. and Wilson, J. R. (2008) 'Understanding the human factors contribution to railway accidents and incidents in Australia', *Accident Analysis and Prevention*, 40(5), pp. 1750–1757. doi: 10.1016/j.aap.2008.06.013.

Bento, C., Cardoso, A. and Dias, G. (2005) *Lecture Notes in Artificial Intelligence: Introduction, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi: 10.1007/11595014_1.

Boluki, S. *et al.* (2020) 'Learnable Bernoulli Dropout for Bayesian Deep Learning', 108. Available at: http://arxiv.org/abs/2002.05155.

Boussabaine, A. H. (1996) 'The use of artificial neural networks in construction management: A review', *Construction Management and Economics*, 14(5), pp. 427–436. doi: 10.1080/014461996373296.

Breiman, L. (2001) 'Statistical modeling: The two cultures', *Statistical Science*, 16(3), pp. 199–215. doi: 10.1214/ss/1009213726.

Brown, K. A., Willis, P. G. and Prussia, G. E. (2000) 'Predicting safe employee behavior in the steel industry: Development and test of a sociotechnical model', *Journal of Operations Management*, 18(4), pp. 445–465. doi: 10.1016/S0272-6963(00)00033-4.

Champa, H. N. and AnandaKumar, D. K. R. (2010) 'Artificial Neural Network for Human Behavior Prediction through Handwriting Analysis', *International Journal of Computer Applications*, 2(2), pp. 36–41. doi: 10.5120/629-878.

Chao, L. C. and Skibniewski, M. J. (1994) 'Estimating Construction Productivity', *Journal of Computing in Civil Engineering*, 8(2), pp. 234–251.

Chen, J. H. and Hsu, S. C. (2007) 'Hybrid ANN-CBR model for disputed change orders in construction projects', *Automation in Construction*, 17(1), pp. 56–64. doi: 10.1016/j.autcon.2007.03.003.

Chen, L. and Yang, X. (2019) 'Using EPPM to Evaluate the Effectiveness of Fear Appeal Messages Across Different Media Outlets to Increase the Intention of Breast Self-Examination Among Chinese Women', *Health Communication*, 34(11), pp. 1369–1376. doi: 10.1080/10410236.2018.1493416.

Chen, Q. and Jin, R. (2013) 'Multilevel Safety Culture and Climate Survey for Assessing New Safety Program', *Journal of Construction Engineering and Management*, 139(7), pp. 805–817. doi: 10.1061/(asce)co.1943-7862.0000659.

Cheng, C. W., Lin, C. C. and Leu, S. Sen (2010) 'Use of association rules to explore cause-effect relationships in occupational accidents in the Taiwan construction industry', *Safety Science*, 48(4), pp. 436–444. doi: 10.1016/j.ssci.2009.12.005.

Cheng, M. Y., Tsai, H. C. and Liu, C. L. (2009) 'Artificial intelligence approaches to achieve strategic control over project cash flows', *Automation in Construction*, 18(4), pp. 386–393. doi: 10.1016/j.autcon.2008.10.005.

Cheung, S. O. *et al.* (2006) 'Predicting project performance through neural networks', *International Journal of Project Management*, 24(3), pp. 207–215. doi: 10.1016/j.ijproman.2005.08.001.

Chi, S., Han, S. and Kim, D. Y. (2013) ‘Relationship between Unsafe Working Conditions and Workers’ Behavior and Impact of Working Conditions on Injury Severity in U.S. Construction Industry’, *Journal of Construction Engineering and Management*, 139(7), pp. 826–838. doi: 10.1061/(asce)co.1943-7862.0000657.

Choudhry, R. M. (2014) ‘Behavior-based safety on construction sites: A case study’, *Accident Analysis and Prevention*, 70, pp. 14–23. doi: 10.1016/j.aap.2014.03.007.

Choudhry, R. M. and Fang, D. (2008) ‘Why operatives engage in unsafe work behavior: Investigating factors on construction sites’, *Safety Science*, 46(4), pp. 566–584. doi: 10.1016/j.ssci.2007.06.027.

Choudhry, R. M., Fang, D. and Ahmed, S. M. (2008) ‘Safety Management in Construction: Best Practices in Hong Kong’, *Journal of Professional Issues in Engineering Education and Practice*, 134(1), pp. 20–32. doi: 10.1061/(ASCE)1052-3928(2008)134:1(20).

Choudhry, R. M., Fang, D. and Mohamed, S. (2007) ‘The nature of safety culture: A survey of the state-of-the-art’, *Safety Science*, 45(10), pp. 993–1012. doi: 10.1016/j.ssci.2006.09.003.

Ciarapica, F. E. and Giacchetta, G. (2009) ‘Classification and prediction of occupational injury risk using soft computing techniques: An Italian study’, *Safety Science*, 47(1), pp. 36–49. doi: 10.1016/j.ssci.2008.01.006.

Clarke, S. (2012) ‘The effect of challenge and hindrance stressors on safety behavior and safety outcomes: A meta-analysis’, *Journal of Occupational Health Psychology*, 17(4), pp. 387–397. doi: 10.1037/a0029817.

Cohen, T. N., Wiegmann, D. A. and Shappell, S. A. (2015) ‘Evaluating the reliability of the human factors analysis and classification system', *Aerospace Medicine and Human Performance*, 86(8), pp. 728–735. doi: 10.3357/AMHP.4218.2015.

Cui, L. *et al.* (2013) 'An integrative model of organizational safety behavior', *Journal of Safety Research*, 45, pp. 37–46. doi: 10.1016/j.jsr.2013.01.001.

Curcuruto, M. and Griffin, M. A. (2018) 'Prosocial and proactive "safety citizenship behaviour" (SCB): The mediating role of affective commitment and psychological ownership', *Safety Science,* 104(January), pp. 29–38. doi: 10.1016/j.ssci.2017.12.010.

El-Gohary, K. M., Aziz, R. F. and Abdel-Khalek, H. A. (2017) 'Engineering Approach Using ANN to Improve and Predict Construction Labor Productivity under Different Influences', *Journal of Construction Engineering and Management*, 143(8), p. 04017045. doi: 10.1061/(asce)co.1943-7862.0001340.

El-Kholy, A. M. (2019) 'Exploring the best ANN model based on four paradigms to predict delay and cost overrun percentages of highway projects', *International Journal of Construction Management*, 0(0), pp. 1–19. doi: 10.1080/15623599.2019.1580001.

El-Kholy, A. M., Tahwia, A. M. and Elsayed, M. M. (2020) 'Prediction of simulated cost contingency for steel reinforcement in building projects: ANN versus regression-based models', *International Journal of Construction Management*, 0(0), pp. 1–15. doi: 10.1080/15623599.2020.1741492.

ElBardissi, A. W. *et al.* (2007) 'Application of the Human Factors Analysis and Classification System Methodology to the Cardiovascular Surgery Operating Room', *Annals of Thoracic Surgery*, 83(4), pp. 1412–1419. doi: 10.1016/j.athoracsur.2006.11.002.

Emsley, M. W. *et al.* (2002) 'Data modelling and the application of a neural network approach to the prediction of total construction costs', *Construction Management and Economics*, 20(6), pp. 465–472. doi: 10.1080/01446190210151050.

Fang, D. P. *et al.* (2004) 'Factor analysis-based studies on construction workplace safety management in China', *International Journal of Project Management*, 22(1), pp. 43–49. doi: 10.1016/S0263-7863(02)00115-1.

Fang, D., Wu, C. and Wu, H. (2015) 'Impact of the Supervisor on Worker Safety Behavior in Construction Projects', *Journal of Management in Engineering*, 31(6), p. 04015001. doi: 10.1061/(asce)me.1943-5479.0000355.

Fang, D., Zhao, C. and Zhang, M. (2016a) 'A Cognitive Model of Construction Workers' Unsafe Behaviors', *Journal of Construction Engineering and Management*, 142(9), p. 04016039. doi: 10.1061/(asce)co.1943-7862.0001118.

Fang, D., Zhao, C. and Zhang, M. (2016b) 'A Cognitive Model of Construction Workers' Unsafe Behaviors', *Journal of Construction Engineering and Management*, 142(9), p. 04016039. doi: 10.1061/(asce)co.1943-7862.0001118.

Fishbein and Ajzen (2010) *Predicting C hanging B ehavior*.

Fishbein, M. and Ajzen, I. (2011) *Predicting C hanging B ehavior*.

Fogarty, G. J. and Shaw, A. (2010) 'Safety climate and the Theory of Planned Behavior: Towards the prediction of unsafe behavior', *Accident Analysis and Prevention*, 42(5), pp. 1455–1459. doi: 10.1016/j.aap.2009.08.008.

Fugas, C. S., Silva, S. A. and Meliá, J. L. (2012) 'Another look at safety climate and safety behavior: Deepening the cognitive and social mediator mechanisms', *Accident Analysis and Prevention*, 45, pp. 468–477. doi: 10.1016/j.aap.2011.08.013.

Garrett, J. W. and Teizer, J. (2009) ‘Human Factors Analysis Classification System Relating to Human Error Awareness Taxonomy in Construction Safety’, *Journal of Construction Engineering and Management*, 135(8), pp. 754–763. doi: 10.1061/(asce)co.1943-7862.0000034.

Goh, Y. M. *et al.* (2018) ‘Factors influencing unsafe behaviors: A supervised learning approach’, *Accident Analysis and Prevention*, 118(June), pp. 77–85. doi: 10.1016/j.aap.2018.06.002.

Goh, Y. M. and Binte Sa’adon, N. F. (2015) ‘Cognitive Factors Influencing Safety Behavior at Height: A Multimethod Exploratory Study’, *Journal of Construction Engineering and Management*, 141(6), p. 04015003. doi: 10.1061/(asce)co.1943-7862.0000972.

Goh, Y. M. and Chua, D. (2013) ‘Neural network analysis of construction safety management systems: a case study in Singapore’, *Construction Management and Economics*, 31(5), pp. 460–470. doi: 10.1080/01446193.2013.797095.

Guo, B. H. W., Goh, Y. M. and Le Xin Wong, K. (2018) ‘A system dynamics view of a behavior-based safety program in the construction industry’, *Safety Science*, 104(January), pp. 202–215. doi: 10.1016/j.ssci.2018.01.014.

Guo, B. H. W., Yiu, T. W. and González, V. A. (2015) ‘Identifying behaviour patterns of construction safety using system archetypes’, *Accident Analysis and Prevention*, 80, pp. 125–141. doi: 10.1016/j.aap.2015.04.008.

Guo, B. H. W., Yiu, T. W. and González, V. A. (2016) ‘Predicting safety behavior in the construction industry: Development and test of an integrative model’, *Safety Science*, 84, pp. 1–11. doi: 10.1016/j.ssci.2015.11.020.

Hamid, A. R. A., Majid, M. Z. A. and Singh, B. (2008) ‘Causes of Accidents

At Construction Sites’, *Malaysian Journal of Civil Engineering*, 20(2), pp. 242–259. Available at: http://myais.fsktm.um.edu.my/6409/1/MJCE09.pdf.

Hofmann, D. A. and Stetzer, A. (1996) ‘A cross-level investigation of factors influencing unsafe behaviors and accidents’, *Personnel Psychology,* 49(2), pp. 307–339. doi: 10.1111/j.1744-6570.1996.tb01802.x.

Hollnagel, E. (2002) ‘Understanding accidents - From root causes to performance variability’, in *IEEE Conference on Human Factors and Power Plants*, pp. 1–6. doi: 10.1109/hfpp.2002.1042821.

Hollnagel, E. (2013) *Resilience Engineering in Practice: A Guidebook* .

Huang, Y. P. *et al.* (2016) ‘Risk perception, risk propensity, and unsafe behavior: An empirical study of workers in Chinese construction industry’, *IEEE International Conference on Industrial Engineering and Engineering Management,* 2016-Decem, pp. 1121–1125. doi: 10.1109/IEEM.2016.7798052.

Ismail, Z., Doostdar, S. and Harun, Z. (2012) ‘Factors influencing the implementation of a safety management system for construction sites’, *Safety Science,* 50(3), pp. 418–423. doi: 10.1016/j.ssci.2011.10.001.

Jemmott, J. B. (2012) ‘The reasoned action approach in HIV risk-reduction strategies for adolescents’, *Annals of the American Academy of Political and Social Science,* 640(1), pp. 150–172. doi: 10.1177/0002716211426096.

Jha, K. N. and Chockalingam, C. T. (2011) ‘Prediction of schedule performance of Indian construction projects using an artificial neural network’, *Construction Management and Economics,* 29(9), pp. 901–911. doi: 10.1080/01446193.2011.608691.

Johari, S. and Jha, K. N. (2020) ‘Interrelationship among Belief, Intention, Attitude, Behavior, and Performance of Construction Workers’, *Journal of Management in Engineering*, 36(6), p. 04020081. doi: 10.1061/(asce)me.1943-5479.0000851.

Johnson, S. E. and Hall, A. (2005) ‘The prediction of safe lifting behavior: An application of the theory of planned behavior’, *Journal of Safety Research*, 36(1), pp. 63–73. doi: 10.1016/j.jsr.2004.12.004.

Juszczyk, M., Leśniak, A. and Zima, K. (2018) ‘ANN Based Approach for Estimation of Construction Costs of Sports Fields’, *Complexity*, 2018. doi: 10.1155/2018/7952434.

Kavzoglu, T. and Mather, P. M. (2003) ‘The use of backpropagating artificial neural networks in land cover classification’, *International Journal of Remote Sensing*, 24(23), pp. 4907–4938. doi: 10.1080/0143116031000114851.

Khan, S. N. (2014) ‘Qualitative Research Method: Grounded Theory’, *International Journal of Business and Management*, 9(11). doi: 10.5539/ijbm.v9n11p224.

Khosravi, Y. *et al.* (2014) ‘Factors influencing unsafe behaviors and accidents on construction sites: A review’, *International Journal of Occupational Safety and Ergonomics*, 20(1), pp. 111–125. doi: 10.1080/10803548.2014.11077023.

Kingma, D. P. and Ba, J. L. (2015) ‘Adam: A method for stochastic optimization’, in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–15.

Kiriakidis, S. (2017) ‘Perceived Behavioural Control in the Theory of Planned Behaviour: Variability of Conceptualization and Operationalization and Implications for Measurement', pp. 197–202. doi: 10.1007/978-3-319-33865-1_25.

Kotu, V. and Deshpande, B. (2019) *Recommendation Engines, Data Science (pp. 343–394)*. doi: 10.1016/b978-0-12-814761-0.00011-3.

Larouzee, J. and Le Coze, J. C. (2020) 'Good and bad reasons: The Swiss cheese model and its critics', *Safety Science*, 126(June 2019), p. 104660. doi: 10.1016/j.ssci.2020.104660.

Liao, C. W. and Perng, Y. H. (2008) 'Data mining for occupational injuries in the Taiwan construction industry', *Safety Science*, 46(7), pp. 1091–1102. doi: 10.1016/j.ssci.2007.04.007.

Lingard, H., Cooke, T. and Blismas, N. (2012) 'Do Perceptions of Supervisors' Safety Responses Mediate the Relationship between Perceptions of the Organizational Safety Climate and Incident Rates in the Construction Supply Chain?', *Journal of Construction Engineering and Management*, 138(2), pp. 234–241. doi: 10.1061/(asce)co.1943-7862.0000372.

Locke, K. (2002) 'The grounded theory approach to qualitative research.', *F. Drasgow & N. Schmitt (Eds.)*, Measuring.

Love, P. E. D. and Smith, J. (2016) 'Error management: Implications for construction', *Construction Innovation*, 16(4), pp. 418–424. doi: 10.1108/CI-01-2016-0001.

Ma, L., Guo, H. and Fang, Y. (2021) 'Analysis of Construction Workers' Safety Behavior Based on Myers-Briggs Type Indicator Personality Test in a Bridge Construction Project', *Journal of Construction Engineering and Management*, 147(1), p. 04020149. doi: 10.1061/(asce)co.1943-7862.0001958.

Man, S. S., Chan, A. H. S. and Wong, H. M. (2017) 'Risk-taking behaviors of Hong Kong construction workers – A thematic study', *Safety Science*, 98, pp. 25–36. doi: 10.1016/j.ssci.2017.05.004.

McDermott, J. H. (2009) 'The cocktail party problem', *Current Biology*, 19(22), pp. 1875–1902. doi: 10.1016/j.cub.2009.09.005.

Mitchell, T. M. (2006) 'The Discipline of Machine Learning', *Machine Learning*, 17(July), pp. 1–7. Available at: http://www-cgi.cs.cmu.edu/~tom/pubs/MachineLearningTR.pdf.

Mitropoulos, P., Abdelhamid, T. S. and Howell, G. A. (2005) 'Systems Model of Construction Accident Causation', *Journal of Construction Engineering and Management*, 131(7), pp. 816–825. doi: 10.1061/(asce)0733-9364(2005)131:7(816).

Mitropoulos, P., Cupido, G. and Namboodiri, M. (2009) 'Cognitive Approach to Construction Safety: Task Demand-Capability Model', *Journal of Construction Engineering and Management*, 135(9), pp. 881–889. doi: 10.1061/(asce)co.1943-7862.0000060.

Mohamed, S. (2002) 'Safety Climate in Construction Site Environments', *Journal of Construction Engineering and Management*, 128(5), pp. 375–384. doi: 10.1061/(ASCE)0733-9364(2002)128:5(375).

Mohamed, S., Ali, T. H. and Tam, W. Y. V. (2009) 'National culture and safe work behaviour of construction workers in Pakistan', *Safety Science*, 47(1), pp. 29–35. doi: 10.1016/j.ssci.2008.01.003.

Mullen, J. (2004) 'Investigating factors that influence individual safety behavior at work', *Journal of Safety Research*, 35(3), pp. 275–285. doi: 10.1016/j.jsr.2004.03.011.

Networks, N. *et al.* (1994) ‘Neural networks in c i v i l e n g i n e e r i n g .’, 8(2).

Nordstrom, S. N. (2015) ‘Not So Innocent Anymore: Making Recording Devices Matter in Qualitative Interviews’, *Qualitative Inquiry*, 21(4), pp. 388–401. doi: 10.1177/1077800414563804.

Patel, D. A. and Jha, K. N. (2015a) ‘Neural Network Approach for Safety Climate Prediction’, *Journal of Management in Engineering*, 31(6), p. 05014027. doi: 10.1061/(asce)me.1943-5479.0000348.

Patel, D. A. and Jha, K. N. (2015b) ‘Neural Network Model for the Prediction of Safe Work Behavior in Construction Projects’, *Journal of Construction Engineering and Management*, 141(1), p. 04014066. doi: 10.1061/(asce)co.1943-7862.0000922.

Patel, D. A. and Jha, K. N. (2016) ‘Evaluation of construction projects based on the safe work behavior of co-employees through a neural network model’, *Safety Science*, 89, pp. 240–248. doi: 10.1016/j.ssci.2016.06.020.

Peško, I. *et al.* (2017) ‘Estimation of costs and durations of construction of urban roads using ANN and SVM’, *Complexity*, 2017. doi: 10.1155/2017/2450370.

Polat, G. (2007) ‘Ann Approach To Determine Cost Contingency’, (1).

Polat, G. *et al.* (2016) ‘Comparison of ANN and MRA Approaches to Estimate Bid Mark-up Size in Public Construction Projects’, *Procedia Engineering*, 164, pp. 331–338. doi: 10.1016/j.proeng.2016.11.627.

Polit, D. F. and Beck, C. T. (2006) ‘The content validity index: Are you sure you know what’s being reported? Critique and recommendations’, *Research in Nursing and Health*, 29(5), pp. 489–497. doi: 10.1002/nur.20147.

Poor Sabet, P. G. (2013) ‘Application of Domino Theory to Justify and Prevent Accident Occurance in Construction Sites’, *IOSR Journal of Mechanical and Civil Engineering*, 6(2), pp. 72–76. doi: 10.9790/1684-0627276.

Poulter, D. R. *et al.* (2008) ‘An application of the theory of planned behaviour to truck driving behaviour and compliance with regulations’, *Accident Analysis and Prevention*, 40(6), pp. 2058–2064. doi: 10.1016/j.aap.2008.09.002.

Ramón, J. P. M. and Rodríguez, F. M. M. (2020) ‘What if violent behavior was a coping strategy? Approaching a model based on artificial neural networks’, *Sustainability (Switzerland)*, 12(18), pp. 1–16. doi: 10.3390/SU12187396.

Reason, J. *et al.* (1990) ‘Errors and violations on the roads: A real distinction?’, *Ergonomics*, 33(10–11), pp. 1315–1332. doi: 10.1080/00140139008925335.

Rebaño-Edwards, S. (2007) ‘Modelling perceptions of building quality-A neural network approach’, *Building and Environment*, 42(7), pp. 2762–2777. doi: 10.1016/j.buildenv.2006.07.018.

Reinach, S. and Viale, A. (2006) ‘Application of a human error framework to conduct train accident/incident investigations’, *Accident Analysis and Prevention*, 38(2), pp. 396–406. doi: 10.1016/j.aap.2005.10.013.

Rosa, J. P. S. *et al.* (2020) ‘Overview of Artificial Neural Networks’, *SpringerBriefs in Applied Sciences and Technology*, pp. 21–44. doi: 10.1007/978-3-030-35743-6_3.

Rostamabadi, A. *et al.* (2019) ‘Model for A Novel Fuzzy Bayesian Network-HFACS (FBN-HFACS) model for analyzing Human and Organizational Factors (HOFs) in process accidents’, *Process Safety and Environmental Protection*, 132, pp. 59–72. doi: 10.1016/j.psep.2019.08.012.

Rutakumwa, R. *et al.* (2020) ‘Conducting in-depth interviews with and without voice recorders: a comparative analysis’, *Qualitative Research*, 20(5), pp. 565–581. doi: 10.1177/1468794119884806.

Sandhya, W. T. and Philominal Judit, L. (2020) ‘Cost Evaluation of Construction Using Artificial Neural Network (Ann)’, *International Research Journal of Modernization in Engineering Technology and Science @International Research Journal of Modernization in Engineering*, 776(04), pp. 2582–5208.

Sato, Y. (1995) ‘Artificial Neural Networks and Statistical Model’, *Japanese journal of applied statistics*, 24(2), pp. 77–88. doi: 10.5023/jappstat.24.77.

Seo, D. C. (2005) ‘An explicative model of unsafe work behavior’, *Safety Science*, 43(3), pp. 187–211. doi: 10.1016/j.ssci.2005.05.001.

Shappell, S. a and Wiegmann, D. a (2000) ‘The Human Factors Analysis and Classification System – HFACS’, *USDOT/FAA/AM-00/7 Office of Aviation Medicine*, p. 19.

Shin, D. P., Gwak, H. S. and Lee, D. E. (2015) *Modeling the predictors of safety behavior in construction workers*, *International Journal of Occupational Safety and Ergonomics*. doi: 10.1080/10803548.2015.1085164.

Shin, M. *et al.* (2014) ‘A system dynamics approach for modeling construction workers’ safety attitudes and behaviors’, *Accident Analysis and Prevention*, 68, pp. 95–105. doi: 10.1016/j.aap.2013.09.019.

Sing, C. P. *et al.* (2014) ‘Personality and Occupational Accidents: Bar Benders in Guangdong Province, Shenzhen, China’, *Journal of Construction Engineering and Management*, 140(7), p. 05014005. doi: 10.1061/(asce)co.1943-7862.0000858.

Singh, K. *et al.* (1995) ‘A theory of reasoned action perspective of voting behavior: Model and empirical test’, *Psychology & Marketing*, 12(1), pp. 37–51. doi: 10.1002/mar.4220120104.

Susmaga, R. (2004) ‘Confusion Matrix Visualization’, in *Intelligent Information Processing and Web Mining*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 107–116. doi: 10.1007/978-3-540-39985-8_12.

Tam, C. M. and Tong, T. K. L. (2003) ‘GA-ANN model for optimizing the locations of tower crane and supply points for high-rise public housing construction’, *Construction Management and Economics*, 21(3), pp. 257–266. doi: 10.1080/0144619032000049665.

Tang, N. *et al.* (2021) ‘A personalized Human Factors Analysis and Classification System (HFACS) for construction safety managementbased on context-aware technology’, *Enterprise Information Systems*, 00(00), pp. 1–26. doi: 10.1080/17517575.2021.1878283.

Terry, D. J. and O’Leary, J. E. (1995) ‘The theory of planned behaviour: The effects of perceived behavioural control and self-efficacy’, *British Journal of Social Psychology*, 34(2), pp. 199–220. doi: 10.1111/j.2044-8309.1995.tb01058.x.

Theophilus, S. C. *et al.* (2017) ‘Human factors analysis and classification system for the oil and gas industry (HFACS-OGI)’, *Reliability Engineering and System Safety*, 167(April 2016), pp. 168–176. doi: 10.1016/j.ress.2017.05.036.

Tixier, A. J. P. *et al.* (2014) ‘Psychological antecedents of risk-taking behavior in construction’, *Journal of Construction Engineering and Management*, 140(11), pp. 1–10. doi: 10.1061/(ASCE)CO.1943-7862.0000894.

Tixier, A. J. P. *et al.* (2016) ‘Application of machine learning to construction injury prediction', *Automation in Construction*, 69, pp. 102–114. doi: 10.1016/j.autcon.2016.05.016.

Tomás, J. M., Meliá, J. L. and Oliver, A. (1999) 'A cross-validation of a structural equation model of accidents: Organizational and psychological variables as predictors of work safety', *Work and Stress*, 13(1), pp. 49–58. doi: 10.1080/026783799296183.

Uğurlu, Ö. *et al.* (2018) 'Modified human factor analysis and classification system for passenger vessel accidents (HFACS-PV)', *Ocean Engineering*, 161(August 2017), pp. 47–61. doi: 10.1016/j.oceaneng.2018.04.086.

Wang, J., Zou, P. X. W. and Li, P. P. (2016) 'Critical factors and paths influencing construction workers' safety risk tolerances', *Accident Analysis and Prevention*, 93, pp. 267–279. doi: 10.1016/j.aap.2015.11.027.

Weber, S. L. *et al.* (1999) 'Model for Construction Budget Performance—Neural Network Approach', *Journal of Construction Engineering and Management*, 125(3), pp. 213–215. doi: 10.1061/(asce)0733-9364(1999)125:3(213).

Wei, C. H. and Lee, Y. (2007) 'Sequential forecast of incident duration using Artificial Neural Network models', *Accident Analysis and Prevention*, 39(5), pp. 944–954. doi: 10.1016/j.aap.2006.12.017.

Wiegand, D. M. (2007) 'Exploring the role of emotional intelligence in behavior-based safety coaching', *Journal of Safety Research*, 38(4), pp. 391–398. doi: 10.1016/j.jsr.2007.03.010.

Wiegmann, D. A. *et al.* (2004) 'Safety culture: An integrative review', *International Journal of Aviation Psychology*. Taylor and Francis Inc., pp. 117–134. doi: 10.1207/s15327108ijap1402_1.

Wiegmann, D. A. and Shappell, S. A. (2001) ‘Human error analysis of commercial aviation accidents: Application of the human factors analysis and classification system (HFACS)’, *Aviation Space and Environmental Medicine*, 72(11), pp. 1006–1016.

Williams, M. and Moser, T. (2019) ‘The Art of Coding and Thematic Exploration in Qualitative Research’, *International Management Review*, 15(1), pp. 45–55.

Wu, W., Gibb, A. G. F. and Li, Q. (2010) ‘Accident precursors and near misses on construction sites: An investigative tool to derive information from accident databases’, *Safety Science*, 48(7), pp. 845–858. doi: 10.1016/j.ssci.2010.04.009.

Wu, Y. chen and Feng, J. wen (2018) ‘Development and Application of Artificial Neural Network’, *Wireless Personal Communications*, 102(2), pp. 1645–1656. doi: 10.1007/s11277-017-5224-x.

Xia, N. *et al.* (2017) ‘Do we see how they perceive risk? An integrated analysis of risk perception and its effect on workplace safety behavior’, *Accident Analysis and Prevention*, 106(June), pp. 234–242. doi: 10.1016/j.aap.2017.06.010.

Xia, N. *et al.* (2018) ‘A hybrid BN-HFACS model for predicting safety performance in construction projects’, *Safety Science*, 101(October 2017), pp. 332–343. doi: 10.1016/j.ssci.2017.09.025.

Xu, S., Zhang, M. and Hou, L. (2019) ‘Formulating a learner model for evaluating construction workers’ learning ability during safety training’, *Safety Science*, 116(August 2018), pp. 97–107. doi: 10.1016/j.ssci.2019.03.002.

Xu, S., Zou, P. X. W. and Luo, H. (2018) ‘Impact of Attitudinal Ambivalence on Safety Behaviour in Construction’, *Advances in Civil Engineering*, 2018. doi:

10.1155/2018/7138930.

Yadav, S. and Shukla, S. (2016) ‘Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification’, in *Proceedings - 6th International Advanced Computing Conference, IACC 2016*. IEEE, pp. 78–83. doi: 10.1109/IACC.2016.25.

Yang, K. *et al.* (2016) ‘Semi-supervised near-miss fall detection for ironworkers with a wearable inertial measurement unit’, *Automation in Construction*, 68, pp. 194–202. doi: 10.1016/j.autcon.2016.04.007.

Ye, G. *et al.* (2018) ‘Improved HFACS on Human Factors of Construction Accidents: A China Perspective’, *Advances in Civil Engineering*, 2018. doi: 10.1155/2018/4398345.

Yıldırım, U., Başar, E. and Uğurlu, Ö. (2019) ‘Assessment of collisions and grounding accidents with human factors analysis and classification system (HFACS) and statistical methods’, *Safety Science*, 119(November 2016), pp. 412–425. doi: 10.1016/j.ssci.2017.09.022.

Yu, Q. Z. *et al.* (2014) ‘Analysis of factors influencing safety management for metro construction in China’, *Accident Analysis and Prevention*, 68, pp. 131–138. doi: 10.1016/j.aap.2013.07.016.

Zaid Alkilani, S. and Jupp, J. (2013) ‘Issues of Construction Health and Safety in Developing Countries: A Case of Jordan’, *The Australasian Journal of Construction Economics and Building*. doi: 10.3316/informit.604959233854855.

Zhang, Cao, T. and Zhao, X. (2019) ‘Using Smartphones to Detect and Identify Construction Workers’ Near-Miss Falls Based on ANN’, *Journal of Construction Engineering and Management*, 145(1), p. 04018120. doi: 10.1061/(asce)co.1943-

7862.0001582.

Zhou, Q., Fang, D. and Mohamed, S. (2011) ‘Safety Climate Improvement: Case Study in a Chinese Construction Company’, *Journal of Construction Engineering and Management*, 137(1), pp. 86–95. doi: 10.1061/(asce)co.1943-7862.0000241.

Zhou, Q., Fang, D. and Wang, X. (2008) ‘A method to identify strategies for the improvement of human safety behavior by considering safety climate and personal experience’, *Safety Science,* 46(10), pp. 1406–1419. doi: 10.1016/j.ssci.2007.10.005.

Zou, P. X. W. and Sunindijo, R. Y. (2013) ‘Skills for managing safety risk, implementing safety task, and developing positive safety climate in construction project’, *Automation in Construction,* 34, pp. 92–100. doi: 10.1016/j.autcon.2012.10.018.

Zupan, J. (1994) ‘Introduction to artificial neural network (ANN) methods: what they are and how to use them’, *Acta Chimica Slovenica*, 41(September), pp. 327–327.