buildings

MDPI

Article

Predicting Construction Workers' Intentions to Engage in Unsafe Behaviours Using Machine Learning Algorithms and Taxonomy of Personality

Yifan Gao [1], Vicente A. González [2,*], Tak Wing Yiu [3], Guillermo Cabrera-Guerrero [4] and Ruiqi Deng [5]

[1] College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China; yfgao91@zju.edu.cn

[2] Civil and Environmental Engineering Department, Faculty of Engineering, University of Alberta, Edmonton, AB T6G 2R3, Canada

[3] School of Built Environment, Massey University, Auckland 4472, New Zealand; tyiu@massey.ac.nz

[4] Escuela de Ingeniería Informática, Pontificia Universidad Católica de Valparaíso, Valparaíso 2950, Chile; guillermo.cabrera@pucv.cl

[5] Department of Educational Technology, Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou 311121, China; r.deng@hznu.edu.cn

* Correspondence: vagonzal@ualberta.ca

Abstract: Dynamic environmental circumstances can sometimes be incompatible with proactive human intentions of being safe, leading individuals to take unintended risks. Behaviour predictions, as performed in previous studies, are found to involve environmental circumstances as predictors, which might thereby result in biased safety conclusions about individuals' inner intentions to engage in unsafe behaviours. This research calls attention to relatively less-understood worker intentions and provides a machine learning (ML) approach to help understand workers' intentions to engage in unsafe behaviours based on the workers' inner drives, i.e., personality. Personality is consistent across circumstances and allows insight into one's intentions. To mathematically develop the approach, data on personality and behavioural intentions was collected from 268 workers. Five ML architectures—backpropagation neural network (BP-NN), decision tree, support vector machine, k-nearest neighbours, and multivariate linear regression—were used to capture the predictive relationship. The results showed that BP-NN outperformed other algorithms, yielding minimal prediction loss, and was determined to be the best approach. The approach can generate quantifiable predictions to understand the extent of workers' inner intentions to engage in unsafe behaviours. Such knowledge is useful for understanding undesirable aspects in different workers in order to recommend suitable preventive strategies for workers with different needs.

Keywords: machine learning; personality configuration; unsafe-behaving intentions

Citation: Gao, Y.; González, V.A.; Yiu, T.W.; Cabrera-Guerrero, G.; Deng, R. Predicting Construction Workers' Intentions to Engage in Unsafe Behaviours Using Machine Learning Algorithms and Taxonomy of Personality. *Buildings* **2022**, *12*, 841. https://doi.org/10.3390/buildings12060841

Academic Editor: Rafiq Muhammad Choudhry

Received: 4 June 2022
Accepted: 14 June 2022
Published: 16 June 2022

1. Introduction

The construction sector plays a significant role in the economy, contributing 13% to the growth in global GDP in 2018 [1]. However, safety issues follow upon such achievements. In recent years, the construction sector has reached record-high accident rates. In the United States, the construction sector caused 1171 non-fatal and fatal injury claims in 2018, which accounted for 18% of all industrial accidents [2]. In the UK, construction was responsible for the largest portion of all occupational injuries in 2018, with an incidence rate of 21% [3]. Although the application of safety management practices seems to have made an impact on the reduction of injury and fatality rates, occupational accidents remain a pervasive issue [4].

Research has found that many occupational accidents are foreseeable, being the result of people's unsafe behaviour from a retrospective point of view [5]. It has been suggested that behaviour prediction could be useful in the identification of vulnerable workers who are risk-takers and in the design of preventive strategies prior to the occurrence of accidents to contribute to the reduction of injury rates [6]. Behaviour predictions, as performed in previous studies, have been found to involve environmental circumstances as predictors [7–12]. However, environmental circumstances have dynamic changes which can sometimes be incompatible with proactive human intentions of being safe and lead individuals to take unintended risks [13]. For example, an individual intends, but fails, to use protective equipment such as a dust mask due to temporary unavailability in the workplace. Thus, existing literature suggests that behaviour predictions involving environmental circumstances might result in biased safety conclusions about individuals' inner intentions to engage in safe or unsafe actions [5], which is known as the intention–behaviour gap [1].

While decreased accuracy in predicting behaviour is expected in a dynamic construction environment [13], prediction based on individuals' intentions may be a more suitable approach to carrying out safety-related assessments in order to reveal people's inner susceptibility to unsafe behaviours [14]. Intention refers to an individual's cognitive stance of exerting physical effort to perform a specific behaviour [15]. Referring to the prevailing behaviour theories, such as the Theory of Planned Behaviour (TPB) [16] and Self-determination Theory (SDT) [17], an individual's intention to engage in a safe or unsafe behaviour is indicative of the individual's "favourable/unfavourable" or "should/should not" evaluation for that behaviour, which reflects an individual's inner susceptibility to engaging in unsafe behaviours and allows more accurate predictions. As stated [18], "in literature related to understanding human behaviour, the concept of intention holds a central place" (p. 153). Since the beginning of the 21st century, researchers have increasingly focused on applying approaches from behavioural ecology in studying intentions [14]. Existing studies, e.g., [14,19–21] have accumulated substantial evidence from humans and other diverse animal species and found that "across a variety of animal taxa (including human race), individuals within populations often display remarkable differences in behavioural intentions that are consistent across time and contexts" [21] (p. 339).

Although the advantages of conducting safety assessments on individuals' inner intentions to engage in unsafe behaviours can be identified, there have been only a few empirical tests in this field, i.e., [1,22]. In addition, the studies [1,22] have been descriptive in nature and provided insufficient insights into the constitution of predictors for workers' behavioural intentions, which prevents these studies from reaching an operative prediction framework. To contribute to the growing research literature in this field, this research aims to:

1. Conduct literature review to investigate the predictors for workers' unsafe behavioural intentions at the individual level;
2. Develop a predictive model of construction workers' intentions to engage in unsafe behaviours based on the predictors identified using machine learning (ML) methods, and test its prediction performance;
3. Gain an enhanced understanding of the weights of each predictor in the prediction practice, as the weights constitute the predictive basis for workers' safety behavioural intentions; and
4. Discuss the predictive model's theoretical and practical implications.

This paper is organised as follows. A review of relevant academic literature to identify the predictors is presented first. The conceptual model section theoretically formulates the modelling framework of workers' intentions to engage in safety-related behaviours and the predictors. The research method specifies the stages involved to mathematically develop the predictive model. Then the data collection is presented. The results involve the training and evaluation of the predictive model, and the insights into the weight of each predictor for predicting workers' intentions. The discussion section interprets the model's theoretical and practical implications. Finally, the conclusions summarise the results, discuss the limitations of this study, and recommend future research directions.

2. Literature Review

Research has found that many occupational accidents are foreseeable, being the result of people's unsafe behaviour from a retrospective point of view [4,23,24]. It is suggested that behaviour prediction can be useful in the identification of vulnerable workers who are risk-takers and in the design of preventive strategies prior to the occurrence of accidents to contribute to the reduction of injury rates [8–11]. A review of the existing literature reveals that current studies have developed forecasting models as a function of environmental circumstances to predict construction workers' potential future behaviours [7–12]. However, researchers have suggested that these predictions might not forecast behaviours well in a practical setting [23–25]. According to the prevailing behaviour theories such as the TPB [16] and SDT [17], a behaviour prediction approach is regarded useful only if individuals' subsequent behaviour is performed as predicted. As meta-analyses have shown [26,27], human safety behaviour is not only driven by individuals' own decision-making processes but is also determined by environmental circumstances. The construction environment has dynamic changes, sometimes unforeseeable, where task- and context-related dynamics exist and can sometimes cause individuals to subsequently fail to act on the behavioural consequences as predicted [7,25].

In particular, task-related dynamics refer to the dynamic changes of environmental circumstances that are immediately related to the accomplishment of work tasks such as workplace real-time supply of safety equipment and personal protective equipment (PPE) (e.g., availability of dust masks and sunscreen) [23]. Context-related dynamics refer to the dynamic changes of environmental circumstances that surround the accomplishment of work tasks such as workplace real-time housekeeping (e.g., tidy housekeeping or electrical power cords laying across walkways) and real-time supervision for safety commitment (e.g., the work shifts of site managers could result in different styles of supervision) [23]. For example, the lack of health and safety management efforts could result in PPE supply–demand imbalances and thereby the emergence of PPE availability–unavailability dynamics [28,29]. Prediction based on PPE availability–unavailability dynamics could be irrelevant to future occurrences (e.g., an individual fails to act as predicted to wear PPE due to workplace unavailability). Thus, these circumstances can be incompatible with proactive human intentions of being safe, leading individuals to take risks that are not factually intended [30,31]. In this regard, research shows that site management practices tend to follow a linear process to investigate why unsafe behaviours happen, lacking a holistic understanding of behavioural causation with workers at the "sharp end" of the system (usually) being incorrectly blamed [32,33]. On this account, research has suggested that the unforeseeable nature of environment circumstances could diminish the accuracy in predicting human behaviour and sometimes result in incorrect safety conclusions about an individual's unsafe-behaviour intentions [13].

The above discussion draws on the relevant academic literature to interpret why previous researchers have suggested that existing behaviour prediction tools might not forecast behaviours well in a practical setting. Behavioural prediction based on individuals' intentions, which is consistent across time and contexts and independent of environment dynamics, seems to be a more suitable approach to carry out safety-related assessments in order to reveal actual unsafe-behaviour potentials.

Investigating the mechanism responsible for consistent individual differences in intentions has been of long-standing interest to the scientific community [34]. Over the past few years, the understanding of human intentions has progressed significantly, and a considerable body of evidence has accumulated documenting that the formation of human intentions has personality origins, e.g., [27,35]. Decades of research involving hundreds of thousands of individuals have revealed two meta dimensions of personality traits—stability (including sub-traits neuroticism, agreeableness, and conscientiousness) and plasticity (including sub-trait extraversion)—which provide a basic model for portraying human personality [35]. Genetic polymorphisms in dopamine and serotonin transmembrane proteins determine individual differences in the density and responsive efficiency of dopamine and serotonin [36]. Dopamine and serotonin levels and signalling efficiencies are found to exert regulatory effects on the magnitude of physiological processes in the brain system such as cognition and pleasure seeking, which set individuals out on different personality trajectories for stability and plasticity and are expressed in behavioural intentions [34,35,37–41]. In previous studies [26,42–46], it has also been documented that subjects with less-efficient dopamine and serotonin systems were observed to consistently be more agreeable and conscientious but less extraverted and neurotic, and exhibited stronger intentions to undertake solicitous behaviours (e.g., caring for colleagues' safety at work) but diminished intentions to take risks (e.g., sensation seeking); conversely, the opposite.

Research corroborates that it is not a single trait but a configuration of traits that constitutes the individually expressed intentions to engage in unsafe behaviours [47]. Each dimension of the meta-traits stability and plasticity is one component of the configuration, and an individual's trait levels in the configuration (i.e., how high or low the individual scores on each trait compared to others) drive quantitatively the scale of the individual's inner desire to take risks and help colleagues [47,48]. For example, extensive empirical evidence has shown that individual variation in trait scores is reflected upon different scales measuring individuals' intentions about behaving safely (e.g., "strongly proactive", "proactive", "neither proactive nor negative", "reactive", or "strongly reactive") [25,49–52]. These studies further pointed out that extraversion and neuroticism are positive regulators, and agreeableness and conscientiousness are negative regulators, in the configuration of individuals' unsafe-behaviour intentions, and that the opposite is true for solicitous behaviours (e.g., caring for colleagues' safety at work). Based on these findings, it has been suggested that the intention to engage in safety-related behaviours at the individual level can be predicted on the basis of quantifying such individual variations in personality configurations [47].

Thus, in this research, the authors develop an approach to predict construction workers' intentions to engage in unsafe behaviours by quantifying workers' individual variations in personality.

3. Research Method

As researchers have pointed out [51], in order to mathematically build a predictive function (to address the research objectives of this study), a conceptual model should be firstly proposed to reflect explicitly its operational architecture to allow numerical simulations of the model to be built (i.e., the input and output matrices). In this research, the conceptual model shapes the modelling framework of construction workers' personality configurations and safety behavioural intentions. The personality configurations and dimensions of safety behavioural intentions of construction workers as the input and output matrices are thereby interpreted. As shown in the literature review, human personality is multidimensional, which can be represented by extraversion, neuroticism, agreeableness, and conscientiousness. Behavioural intentions are found to be involved in the decision-making processes for safety compliance and safety participation, which represent the basic dimensions that constitute people's safety performance in the workplace [53]. Safety compliance refers to compliance-related behaviour that individuals carry out to keep themselves safe, such as not taking shortcuts, using safety equipment, and following safety procedures and rules. Safety participation refers to solicitous safety activities such as reporting co-workers' safety problems, keeping the workplace clean, and caring for colleagues' safety, which may not directly contribute to one's own safety but helps to develop an environment that supports safety.

A conceptual model is proposed to represent the modelling framework of the personality configuration and behavioural intention dimensions, as ascertained. As shown in Figure 1a, the conceptual model holds a multiple-input-multiple-output architecture. The input consists of an nx4 matrix, which represents n samples of four personality traits (i.e., extraversion, agreeableness, conscientiousness, and neuroticism). The output is a nx2 matrix, which characterises n samples of two behavioural intention dimensions, namely, safety compliance and participation intentions. The conceptual model was then sketched in MATLAB using the Simulink Model function (see Figure 1b), which provides an abstract description of the mathematical simulation to be performed. After gathering the digital data on construction workers' personality and safety behavioural intentions, the abstract description in the Simulink Model function will be able to absorb digital data as its input and output signals to mathematically build the predictive framework.

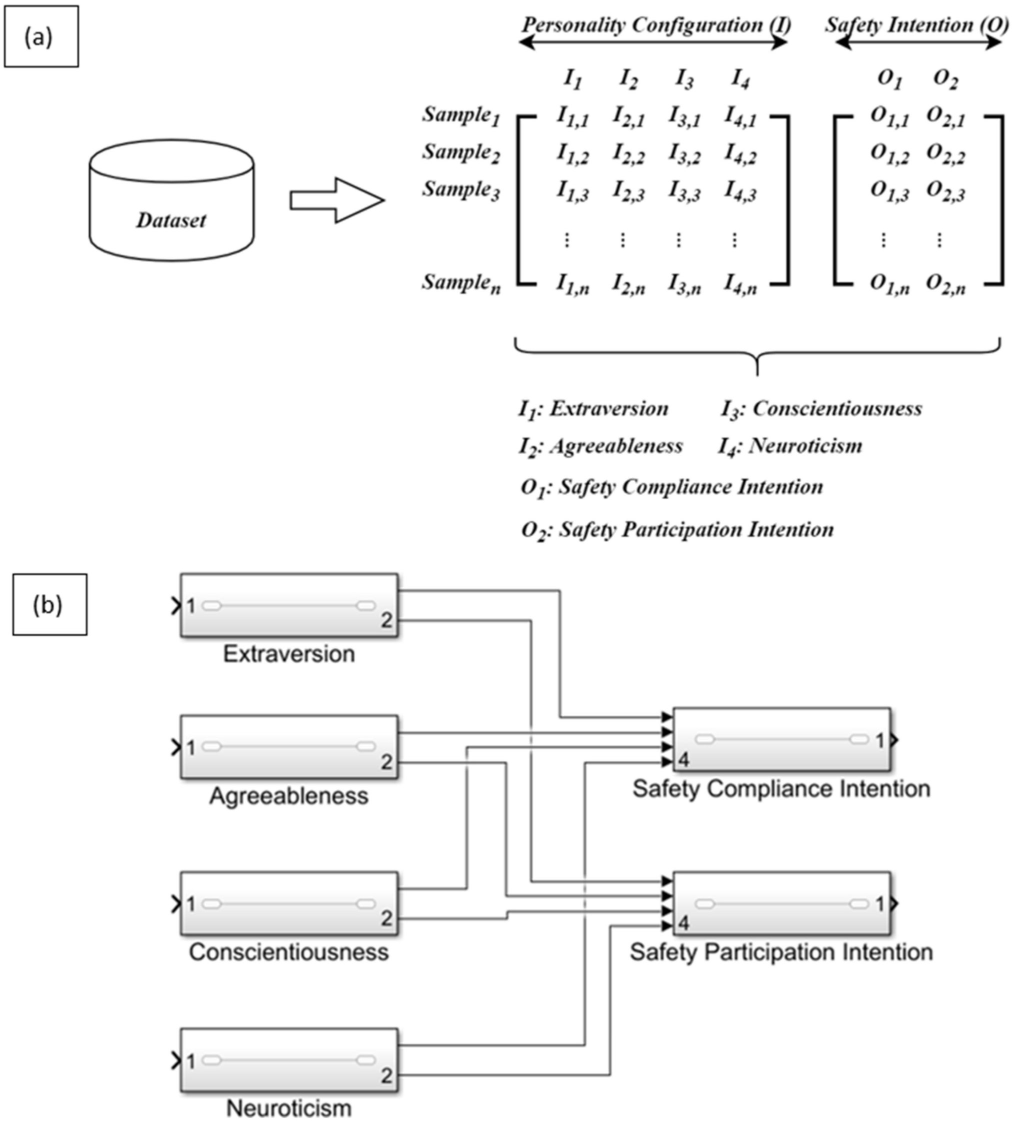

Figure 1. (**a**) Conceptual model; (**b**) Simulink model.

To achieve the objectives of this research, numerical simulations of the modelling framework were firstly carried out to establish the model's mathematical framework. To perform the numerical simulations, the gathering of relevant data concerning construction workers' personality configurations and behavioural intentions is one critical step in this research and is discussed in the data collection section. To further 'crack the code' of the conceptual model into its mathematical equivalent, ML algorithms were applied. In recent years, ML has seen an exponential rise in its usage due to its outstanding ability to iteratively learn from data to look into hidden insights and evolve its architecture to accommodate new findings [54]. Since there are a variety of ML algorithms that can perform data mining tasks, this research experimented with different ML algorithms and evaluated which algorithm is more accurate to capture the predictive relationship between workers' personality configurations and behavioural intentions, to establish a robust and reliable predictive model. Specifically, five ML algorithms—backpropagation neural network (BP-NN), decision tree (DT), support vector machine (SVM), k-nearest neighbours (KNN), and multivariate linear regression (MLR)—were utilised in this research (Figure 2).

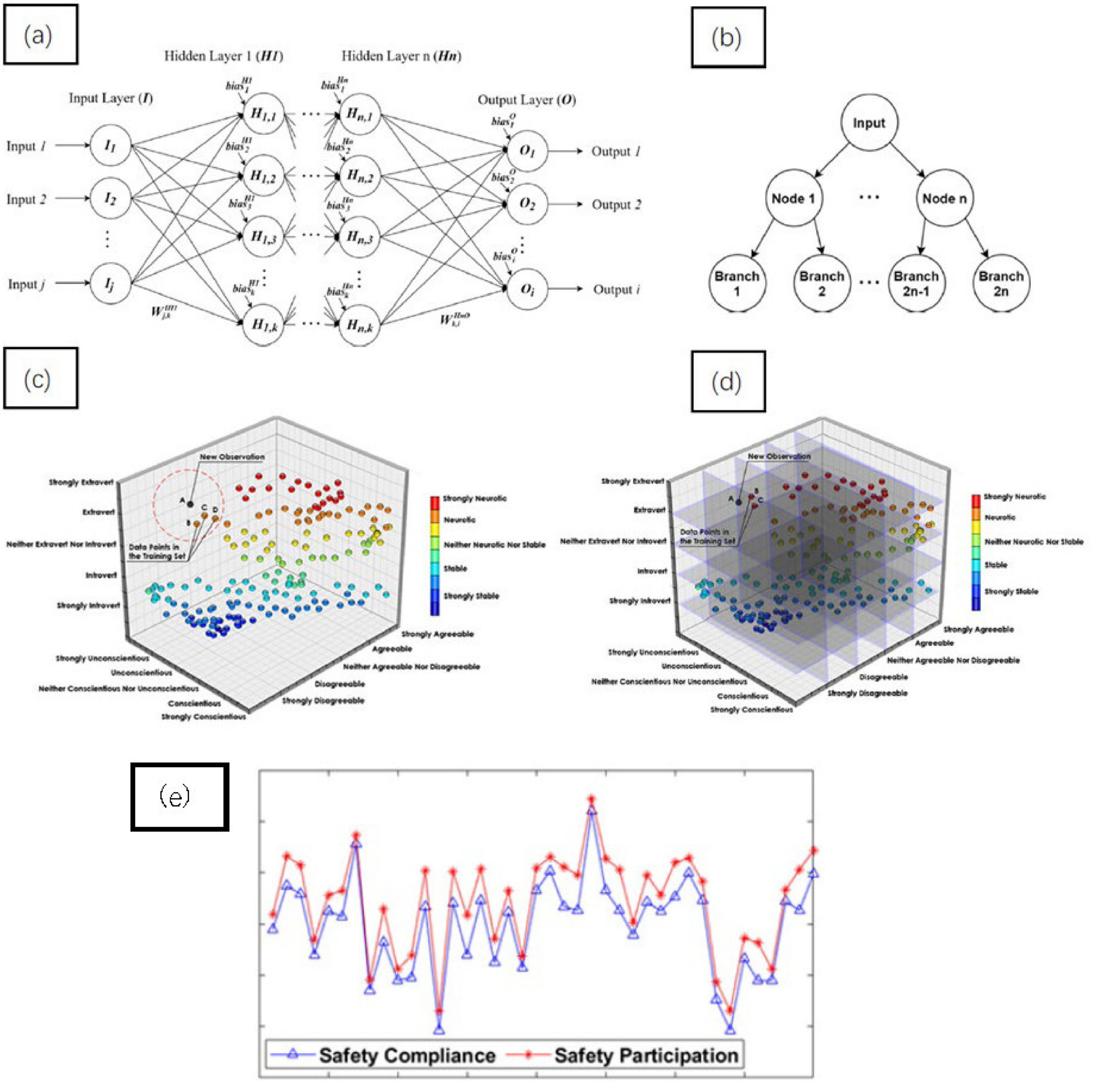

Figure 2. The architecture of utilised machine learning algorithms: (**a**) neural network; (**b**) decision tree; (**c**) support vector machine; (**d**) K-nearest neighbours; (**e**) multivariate linear regression.

A BP-NN consists of multiple layers, where the neurons in the input, hidden, and output layers are responsible, respectively, for receiving input signals, processing the received signals, and decoding the processed signals into the outputs [55]. During a training process, the BP-NN iteratively adjusts the weights (i.e., connection strength between neurons in the input, hidden, and output layers) to minimise the squared difference between the BP-NN estimated values and the actual values in the training dataset [56]. Once trained, the BP-NN is able to obtain outputs for a given set of inputs. DT imitates a tree structure, where binary recursive partitioning is performed to identify the best criteria to divide the standardised inputs into terminal branches such that the squared difference between the values in the terminal branches and the corresponding outputs in the training dataset can be minimised [6]. In KNN, a new observation is predicted based on its similarity to the features of certain existing data points in the training set [6]. KNN analyses the Euclidean distance between a new observation and all the existing data points in the training set and assigns the new observation to the respective class that indicates the shortest distance, as well as the most similar features [54]. SVM identifies optimal separating hyperplanes to break off an n-dimensional space into classes such that a new observation can fall into the rightful boundaries for decision making [55]. MLR indicates that the sole output to be predicted is a linear function of one or more explanatory variables [57]: $y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_ix_i$.

The authors acknowledge that these five methods (BP-NN, DT, SVM, KNN, and MLR) are chronologically not the most up-to-date techniques, as they were firstly published in 1970 [58], 1986 [59], 1992 [60], 1992 [61], and 1979 [62], respectively. There could be machine learning methods that are chronologically relatively new compared to the five methods used in this research. For example, the convolutional neural network (CNN), which represents a major breakthrough in computer vision, has propelled the deep learning field [63]. However, as suggested by data scientists [63], in data analytics, choosing a suitable algorithm for a particular use case (i.e., data type) is an essential task. A CNN approach works well with data containing spatial features and is mainly used for image processing, classification and segmentation [63]. This is because neighbouring pixels of image data indicate spatial features, which are represented in a two-dimensional array format and are thus suitable for feeding into the architecture of a CNN [63]. In this research, the data gathered on construction workers' personality and behavioural intentions is in a tabular format, which does not form a spatial pattern linking different subjects in the table; thus, a CNN is not ideal. An extensive review of the existing literature reveals that the five methods (BP-NN, DT, SVM, KNN, and MLR) are widely used in today's research to generate predictions from tabular data. Supporting examples identified in the literature are from 2018, 2020, and 2021, including technical papers from high-impact journals such as the Journal of Construction Engineering and Management [6], Automation in Construction [55], Advanced Engineering Informatics [64], and Safety Science [65,66]. The validity of BP-NN, DT, SVM, KNN, and MLR has been well-scrutinised not only by their original authors, as mentioned above, but also by other researchers who have employed these methods, e.g., [55,64–66], and they have been proven highly valid and robust for tabular data particular use.

BP-NN, DT, SVM, and KNN are data-driven methods that allow for identifying natural patterns between variables without assuming any preconception in terms of the mathematical structure of the data [55]. MLR is the rule-driven method in ML which derives the relationships between variables by imposing a linearity assumption on the relationships [57]. Although MLR is one of the simplest statistical methods for developing predictive models and only supports linear solutions, it was utilised in this research in order to evaluate the performance of a linear method against the other non-linear ones given the following considerations:

MLR can sometimes outperform data-driven methods such as BP-NN when the underlying relationship is closer to a linear fashion and a smaller number of predictors is involved [67].

The personality–intention relationship has been observed in some cases to be linear [1,52], and the number of personality traits included as the predictors (i.e., extraversion, neuroticism, agreeableness, and conscientiousness) is relatively smaller.

A greater model performance (i.e., minimal prediction loss) usually implies that a more accurate mapping of the authentic relationships between variables has been captured [67]. In order to explore which algorithm can best model the predictive relationship between workers' personality configurations and behavioural intentions, the prediction loss of the BP-NN, DT, SVM, KNN, and MLR models was evaluated using loss functions including mean squared error (MSE), normalized root mean square error (NRMSE), and mean absolute percentage error (MAPE) [68] (see the results section). To gain further insights into the weight of each trait in the personality configurations for workers' behavioural intentions, weight analysis was conducted on the ML model that indicated the best prediction performance, which is discussed in greater detail in the results section. The theoretical and practical implications of the predictive model are interpreted in the discussion section.

4. Data Collection

Collecting relevant data for numerical simulations is a critical step, which was conducted with the objective of gathering data concerning the conceptual model in terms of inputs (personality traits) and outputs (behavioural intention dimensions) from workers. The following subsections discuss in detail the determination of the instruments for data collection on workers' personality traits and safety behavioural intentions as well as the demographic information of the participants.

4.1. Personality Traits

Personality has scales and is hierarchically measurable via its dimensions, including extraversion, conscientiousness, agreeableness, and neuroticism [27]. The existing scales measure personality dimensions along a five-category Likert scale continuum [27] (Table 1).

Table 1. Scales of personality dimensions.

	Extraversion	Conscientiousness	Agreeableness	Neuroticism
Scale 1	Strongly Extravert	Strongly Conscientious	Strongly Agreeable	Strongly Neurotic
Scale 2	Extravert	Conscientious	Agreeable	Neurotic
Scale 3	Neither Extravert Nor Introvert	Neither Conscientious Nor Unconscientious	Neither Agreeable Nor Disagreeable	Neither Neurotic Nor Stable
Scale 4	Introvert	Unconscientious	Disagreeable	Stable
Scale 5	Strongly Introvert	Strongly Unconscientious	Strongly Disagreeable	Strongly Stable

The latest research trends indicate that personality is currently being assessed through explicit techniques [51]. Explicit techniques refer to self-ratings, where individuals are explicitly asked to report their or others' personality characteristics using a number of measurable items in response to certain situations (e.g., "I see myself as someone who tends to be quiet"; on a five-point Likert scale with anchors ranging from "strongly disagree" to "strongly agree") [69]. As researchers have pointed out [51,70], the easy-to-administer personality items used in explicit measures could offer an optimal balance between diagnostic accuracy and feasibility. The explicit instruments—the Big Five Inventory (BFI) [71], NEO Five-Factor Inventory (NEO-FFI) [72], Revised NEO Personality Inventory (NEO-PI-R) [73], and Frame-of-Reference Big Five Inventory (FOR-BFI) [74]—were introduced in the late 80s and early 90s and are to date the most widely used measures of personality [69]. These instruments are easy to administer and cost-effective, and they are not merely screening tools but are built on relevant personality and psychometric theories that were ascertained through decades of research and testing with individuals [75,76]. The psychometric properties (i.e., reliability and validity) of these instruments have been well-scrutinised not only by their original authors, as mentioned above, but also by numerous researchers who have employed these instruments afterwards [51,77–80], proving to be highly reliable and valid instruments.

Among the above four prevalent instruments, BFI, as well as NEO-FFI and NEO-PI-R, were developed purposely for a non-contextualised measure of personality [76,81]. A non-contextualised assessment is an assessment that examines personality without reference to any particular context [82]. For instance, the conscientiousness items in BFI, which include items such as "I see myself as someone who can be somewhat careless" and "I see myself as someone who tends to be disorganised", are not contextually defined and are open to interpretation across a range of contexts (e.g., at work, at home, at school) [83]. In recent years, psychology researchers have reached the consensus that non-contextualised personality items are mostly useful in predicting non-context-dependent criteria (e.g., internet addiction disorder, conduct disorder) but suboptimal for behaviours or intentions that take place under certain contexts (e.g., intentions to engage in unsafe behaviours in the workplace) [80,82–86]. Such consensus reached in the literature also suggests that "specifying situational contexts will likely reduce the range of expectations given a person's dispositional tendency, leading to more accurate predictions in linking personality to behaviour or intention in a certain context" [84] (p. 302). The improvement of predictive validity by means of contextualisation is referred to as the frame-of-reference (FOR) effect (viz., FOR-BFI instrument, as mentioned) [83]. Invented by [74], the FOR notion is to reference a personality instrument to certain contexts via the appendage of contextual references (e.g., "at work", "at home", or "at school") to each item in the instrument. As research suggests [82], "simply by asking participants to complete personality scales with a particular context in mind, the predictive validity of the consequent personality models can be significantly improved" (p. 153). For example, a meta-analysis [87] and recent psychology research [82,88] found that instruments composed of contextualised personality items were able to explain at least two times more variance in individuals' job performance than instruments that did not specify a context.

Taking the above considerations into account, FOR-BFI (in a work context) was used as the instrument for collecting personality data in this research. FOR-BFI (in a work context) measures the dimensions extraversion, agreeableness, conscientiousness, and neuroticism of human personality along a five-point Likert scale (from 1 = disagree strongly to 5 = agree strongly), which is available from https://forms.gle/eREQZH1UpXQH2uFDA (accessed on 16 May 2022).

4.2. Safety Behavioural Intentions

As ascertained in the conceptual model section, the safety behavioural intention we aim to evaluate in this paper consists of two components: safety compliance and safety participation. An extensive review of the existing literature reveals that there are several instruments—namely the Safety Behaviour Scale (SBS) [89], Safety Diagnosis Questionnaire (SDQ) [90], Safety Climate Index: Workers' Health and Safety Behaviour (SCI-WHSB) [91], Work Safety Scale (WSS) [92], and Safety Performance Scale (SPS) [93]—which appear to be popular and have been widely used by researchers in occupational safety domains (e.g., construction, manufacturing) over the years, e.g., [51,70,94–98]. The reliability and validity of these instruments have been well-scrutinised not only by their original authors, as mentioned above, but also by other researchers, e.g., [51,94–98] who have employed these instruments and have been proven to be highly reliable and valid. However, it has been found that some of the instruments (i.e., SDQ, SCI-WHSB, and WSS) fail to provide a thorough measure of behavioural intention and lack the evaluation of safety participation, whereas SBS and SPS have focused on both the safety compliance and safety participation dimensions of behavioural intention.

SBS and SPS were designed to cover the measurement of the sub-dimensions of safety compliance and participation [89,93], namely, not taking shortcuts, using safety equipment, following safety procedures and rules, reporting safety problems, keeping the workplace clean, and caring for colleagues' safety. However, it has been found that SPS contains six items that are all positively worded (e.g., "I use personal protective equipment for the task I am doing") [93]. Using purely positive-worded items in an instrument is considered imbalanced and has the potential to induce positive bias [89]. Meanwhile, in SBS, the majority of statements (i.e., 10 out of 11 items) are framed using both positive and negative items [89] (which is available from https://forms.gle/Ht65oKjXyhZh9cuC9 (accessed on 16 May 2022)):

1. Safety Compliance: Q1 and Q2, Q3 and Q8, Q4 (Q7) and Q9 (from the weblink above).
2. Safety Participation: Q6 (Q11) and Q10 (from the weblink above).

In addition, SBS is a concise instrument (11 items) and at the same time contains adequate contextualised items as needed to measure human intentions to engage in safety-related behaviours in the work context [94]. Taking the above considerations into account, SBS was chosen for use in this research to collect data on construction workers' safety behavioural intentions.

4.3. Participants and Demographic Information

Ethical approval was granted by the University of Auckland Human Participants Ethics Committee. In June 2019, the authors visited six construction projects in Auckland, New Zealand, and asked frontline workers encountered if they were willing to respond to an online survey, which was self-administered and anonymous. Gatekeeper permission was obtained from managers prior to approaching on-site workers. To avoid interference with the construction progress, workers who agreed to the request were provided with a web link so that they could respond to the online survey during off-duty hours. Two hundred and eighty workers finally provided their responses. Twelve out of the 280 responses were deemed unusable due to having unanswered items and dropped from this study. Table 2 reports the distribution of demographic characteristics of the respondents. In responding to self-referential statements, individuals are led to recall relevant episodes from their memory system [94]. It is therefore of great importance that such episodes appear to exist in, or at least make sense to, the participants, which has been found to be determined, in most cases, by individuals' personal experience in the context that is measured (e.g., at work). To understand whether the participants had adequate on-site experience such that they were able to provide useful responses, information on their years of work experience and types of trade was requested for self-reporting. As seen (Table 2), the participants were composed of individuals employed in different construction trades, and more than half of them had worked for a minimum of five years on construction projects. It thus appeared that the participants had adequate on-site experience to generate useful responses.

Table 2. Demographic characteristics.

Features	Categories	Respondents	
		Frequency	Percentile (%)
Age Range	20–29	164	61.2
	30–39	75	28.0
	40–49	27	10.1
	>50	2	0.7
Work Experience (Years)	0–5	132	49.3
	6–10	45	16.7
	11–15	34	12.7
	16–20	36	13.4
	>20	21	7.9
Type of Trade	Mason	44	16.4
	Carpenter and Joiner	20	7.5
	Electrician	13	4.9
	Foreman	11	4.1
	Miscellaneous Labourer	84	31.3
	Plant Operator	25	9.3
	Plumber	24	9.0
	Welder	25	9.3
	Project Manager	15	5.6
	Painter	7	2.6

5. Results

5.1. Data Split Ratio

As specified in the research method section, five ML algorithms—BP-NN, DT, SVM, KNN, and MLR—were adopted to carry out the numerical simulations of the conceptual model proposed. To ascertain which algorithm best modelled the predictive relationship, the performance (i.e., prediction loss) of the five ML algorithms was evaluated and compared. As such, the whole dataset was split into three groups: the training, validation, and testing datasets. The training dataset is used to iteratively adjust the modelling parameters (e.g., the weights between input and hidden neurons of BP-NN) in the ML algorithms such that the algorithms could be optimised to best map the input–output modelling framework and produce minimal prediction losses [55]. The validation dataset was used to determine when training should terminate in order to prevent overfitting [56]. The testing dataset was used to test the eventual prediction losses of the ML algorithms after the training process [56].

The criteria for determining the dataset split ratio have been recommend in the literature [99,100]: (1) The training dataset should be more than two-thirds of the whole dataset and the validation and testing datasets should be equally split among the rest. (2) Each of the validation and testing datasets should be approximately one-fourth to one-eighth of the training dataset. Given the criteria, the commonly applied split ratios include 70:15:15 and 80:10:10 for training, validation and testing datasets [100]. Additionally, the 70:15:15 has been reported to be a more balanced split ratio than 80:10:10, as it includes as many data points for training as possible and also preserves sufficient data portions for validation and testing [99]. Thus, the 70:15:15 ratio was selected for use in this research. The whole dataset had a total of 268 valid samples, of which 188 (70%), 40 (15%), and 40 (15%) samples were randomly assigned for training, validating, and testing of the five ML algorithms.

5.2. Training Process

Training of the five ML algorithms was performed using MATLAB software. During the training process, iterative optimisation of the five ML algorithms was performed to seek the minimal differences between the actual values (i.e., self-reported intentions to engage in unsafe behaviours from samples included in the training set) and the predictions from the algorithms. The functions for difference minimisation are referred to as loss functions (i.e., MSE, NRMSE, and MAPE), where a lower MSE, NRMSE, or MAPE value indicates a decreased actual–algorithm prediction gap as well as a reduced prediction error [68]. According to the literature [101], MSE, NRMSE, and MAPE values lower than 0.25, 0.1, and 0.1, respectively, are the thresholds that a satisfactory error minimisation must achieve. The equations of the loss functions are as follows:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{1}$$

$$\mathrm{NRMSE} = \frac{\mathrm{MSE}}{ymin_{max}} \tag{2}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{3}$$

where n = the number of samples in a dataset; y_i = the actual value for the ith sample; $\hat{y}_i$ = the algorithm estimated value for the ith sample; $y_{\max}$ = the maximum actual value in the dataset; and $y_{\min}$ = the minimum actual value in the dataset.

Through the training process, the prediction error was iteratively minimised and resulted in the MSE, NRMSE, and MAPE values shown in Table 3. BP-NN outperformed the other algorithms and yielded the lowest errors on the training dataset. According to the mentioned threshold (MSE < 0.25, NRMSE < 0.1 and MAPE < 0.1), the optimisation performance of BP-NN, DT, SVM, and KNN was found to be highly satisfactory, whereas MLR revealed a less-than-satisfactory performance.

Table 3. MSE, NRMSE, and MAPE for algorithm optimisation on the training dataset.

ML Algorithms	Safety Compliance Intention			Safety Participation Intention		
	MSE	NRMSE	MAPE	MSE	NRMSE	MAPE
BP-NN	0.042	0.010	0.021	0.073	0.022	0.027
DT	0.051	0.013	0.028	0.103	0.055	0.076
SVM	0.113	0.043	0.047	0.107	0.037	0.047
KNN	0.067	0.020	0.030	0.075	0.023	0.027
MLR	0.450	0.162	0.174	0.634	0.189	0.213

Training to determine the parameters of BP-NN, DT, SVM, KNN, and MLR is interpreted in the following sub-sections.

5.2.1. BP-NN

In this research, the BP-NN consists of four input neurons (one corresponds to each personality trait) and two output neurons (one corresponds to each behavioural intention dimension). To determine the number of hidden neurons, a method widely recommended and used in the literature was followed: $m \leq n \leq 2m$, where n is the number of hidden neurons and m is the number of input neurons, e.g., [102–104]. The BP-NN was thereby trained repeatedly with varying numbers of hidden neurons (from four to eight) and hidden layers to identify the number of hidden neurons and hidden layers that can demonstrate the optimal performance on prediction loss (i.e., minimal MSE). As shown in Table 4, one hidden layer with eight hidden neurons exhibited the optimal performance among its alternatives on the loss functions (i.e., minimal MSE, NRMSE, and MAPE for safety compliance and safety participation predictions), and further addition to the number of hidden layers and neurons appeared to be no longer able to lead to improved BP-NN performance.

Table 4. MSE, NRMSE, and MAPE results on BP-NN training with varying numbers of hidden neurons and layers.

Results on Safety Compliance						
Number of Hidden Neurons	Number of Hidden Layers = 1			Number of Hidden Layers = 2		
	MSE	NRMSE	MAPE	MSE	NRMSE	MAPE
4	0.203	0.051	0.123	0.453	0.113	0.228
5	0.234	0.059	0.142	0.342	0.086	0.181
6	0.176	0.044	0.088	0.311	0.078	0.156
7	0.099	0.025	0.051	0.297	0.081	0.149
8	0.042	0.010	0.021	0.125	0.032	0.065
Results on Safety Participation						
Number of Hidden Neurons	Number of Hidden Layers = 1			Number of Hidden Layers = 2		
	MSE	NRMSE	MAPE	MSE	NRMSE	MAPE
4	0.257	0.064	0.142	0.551	0.137	0.286
5	0.192	0.048	0.132	0.367	0.092	0.184
6	0.163	0.041	0.082	0.324	0.085	0.162
7	0.103	0.026	0.052	0.285	0.073	0.143
8	0.073	0.022	0.027	0.112	0.025	0.051

The computation inside a neuron is carried out by means of an activation function, which determines the output of that neuron, given a set of inputs [54]. As suggested in the literature [105], rectified linear unit (ReLU), SoftPlus, tansig, logsig, and purelin are current widely used activation functions. The BP-NN was trained repeatedly with varying combinations of these activation functions in the hidden and output layers to identify the combination that demonstrated the optimal performance on the loss functions MSE, NRMSE, and MAPE. It was found that the use of tansig and ReLU in the hidden and output layers, respectively, exhibited the optimal performance for safety compliance and safety participation predictions (see Table 5), which was thereby determined for the BP-NN.

Table 5. MSE, NRMSE, and MAPE results on BP-NN training with varying combinations of activation functions in hidden and output layers.

Activation Functions		Results on Safety Compliance			Results on Safety Participation		
Hidden Layer	**Output Layer**	**MSE**	**NRMSE**	**MAPE**	**MSE**	**NRMSE**	**MAPE**
ReLU	SoftPlus	0.132	0.035	0.076	0.251	0.062	0.125
ReLU	tansig	0.651	0.162	0.325	0.723	0.180	0.361
ReLU	logsig	0.442	0.110	0.221	0.571	0.142	0.285
ReLU	purelin	0.087	0.022	0.044	0.096	0.024	0.048
SoftPlus	ReLU	0.424	0.106	0.212	0.571	0.142	0.285
SoftPlus	tansig	0.871	0.218	0.436	0.956	0.239	0.478
SoftPlus	logsig	0.654	0.164	0.327	0.590	0.147	0.295
SoftPlus	purelin	0.245	0.061	0.123	0.329	0.082	0.164
tansig	ReLU	0.042	0.010	0.021	0.073	0.022	0.027
tansig	SoftPlus	0.134	0.034	0.067	0.189	0.047	0.094
tansig	logsig	0.465	0.116	0.232	0.467	0.116	0.233
tansig	purelin	0.589	0.147	0.295	0.674	0.168	0.337
logsig	ReLU	0.243	0.061	0.122	0.386	0.096	0.193
logsig	SoftPlus	0.351	0.088	0.175	0.467	0.116	0.233
logsig	tansig	0.089	0.022	0.044	0.057	0.014	0.028
logsig	purelin	0.167	0.041	0.083	0.189	0.047	0.094
purelin	ReLU	0.092	0.023	0.046	0.089	0.022	0.044
purelin	SoftPlus	0.632	0.158	0.316	0.789	0.197	0.394
purelin	tansig	0.562	0.140	0.281	0.684	0.171	0.342
purelin	logsig	0.367	0.091	0.184	0.479	0.119	0.239

The computation inside the three-layer BP-NN (i.e., one input layer, one hidden layer, one output layer) can therefore be expressed using Equations (4)–(7) below.

$$H_k = \sum_{j=1}^{N_j} w_{j,k}^{IH} I_j + bias_k^H \tag{4}$$

$$H_k' = f_H(H_k) = tansig(H_k) = \frac{2}{1+e^{-2(H_k)}} - 1 \tag{5}$$

$$O_i = \sum_{k=1}^{N_k} w_{k,i}^{HO} H_k' + bias_i^O \tag{6}$$

$$O_i' = f_O(O_i) = ReLU(O_i) = \begin{cases} 0 & \text{for } O_i \le 0 \\ O_i & \text{for } O_i > 0 \end{cases} \tag{7}$$

where H_k = input to the kth hidden neuron; I_j = input to the kth hidden neuron from the jth input neuron (the value of I_j is an individual's score on each personality dimension); $w_{j,k}^{IH}$ = weight between the jth input neuron and kth hidden neuron (the values are provided in Table 6); N_j = the number of input neurons (N_j = 4); $bias_k^H$ = bias of the kth hidden neuron (the values are provided in Table 6); H_k' = output of the kth hidden neuron; $f_H(x)$ = activation function for the hidden layer (i.e., tansig); O_i = input to the ith output neuron; $w_{k,i}^{HO}$ = weight between the kth hidden neuron and ith output neuron (the values are provided in Table 6); N_k = the total number of hidden neurons (N_k = 8); $bias_i^O$ = bias of the ith output neuron (the values are provided in Table 6); O_i' = output of the ith output neuron (i = 1, O_i' = the prediction for safety compliance intention; i = 2, O_i' = the prediction for safety participation intention); and $f_O(x)$ = activation function for the output layer (i.e., ReLU).

Table 6. BP-NN weights and biases.

Input/Output Neurons (I_j/O_i)	Weights/Bias	Hidden Neurons (H_k)							
		H_1	H_2	H_3	H_4	H_5	H_6	H_7	H_8
I_1	$w_{1,k}^{IH}$	0.617	0.100	−0.459	0.065	3.925	1.434	3.006	−2.514
I_2	$w_{2,k}^{IH}$	−4.368	1.439	−2.976	0.342	3.888	−0.671	1.456	1.332
I_3	$w_{3,k}^{IH}$	1.888	−1.036	2.561	−0.369	−0.734	−8.384	−1.224	1.200
I_4	$w_{4,k}^{IH}$	2.222	−0.722	1.070	−0.186	−1.412	−0.783	3.461	0.957
	$bias_k^H$	−1.303	−1.255	2.168	−0.136	0.190	3.776	3.004	−5.919
O_1	$w_{k,1}^{HO}$	−0.469	2.493	0.325	−5.334	0.200	1.535	0.038	1.301
O_2	$w_{k,2}^{HO}$	−0.717	−0.363	−1.189	−4.906	0.284	4.369	−0.826	3.982
	$bias_i^O$	0.774 (for O_1)				0.513 (for O_2)			

5.2.2. DT

During the training process, the following parameters of DT were tuned in order to best fit the training dataset: max_depth, min_sample_split [106]. The parameter max_depth indicates how deep a DT can be [106]. The parameter min_sample_split represents the minimum percentage of samples required to split an internal node [106]. As presented in Figure 3, the training results revealed that (1) for safety compliance prediction, max_depth = 58 and min_sample_split = 45% indicated the optimal performance among its alternatives on the loss functions MSE = 0.051, NRMSE = 0.013, and MAPE = 0.018; (2) for safety participation prediction, max_depth = 64 and min_sample_split = 48% indicated the optimal performance among its alternatives on the loss functions MSE = 0.103, NRMSE = 0.055, and MAPE = 0.076. The above values of max_depth and min_sample_split were therefore determined for the DT model in this research.

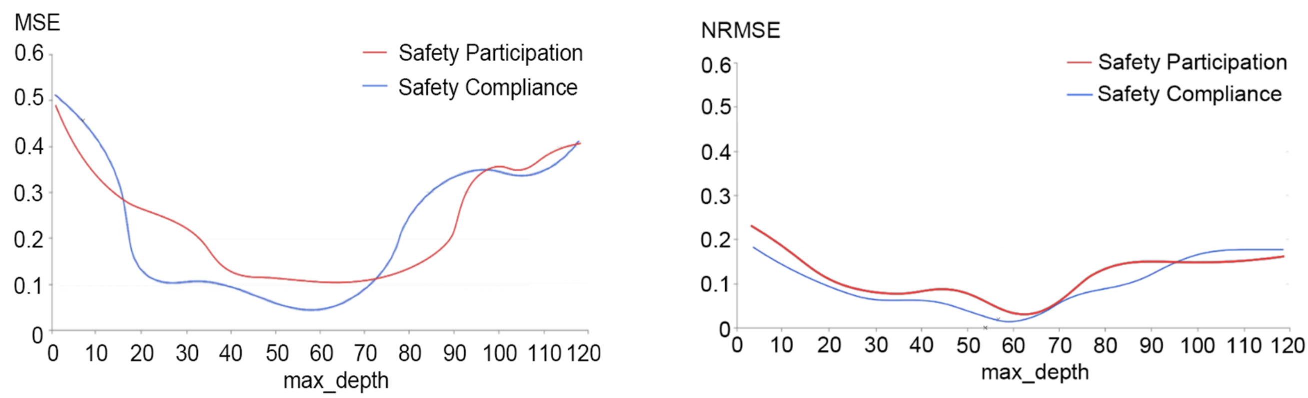

Figure 3. *Cont.*

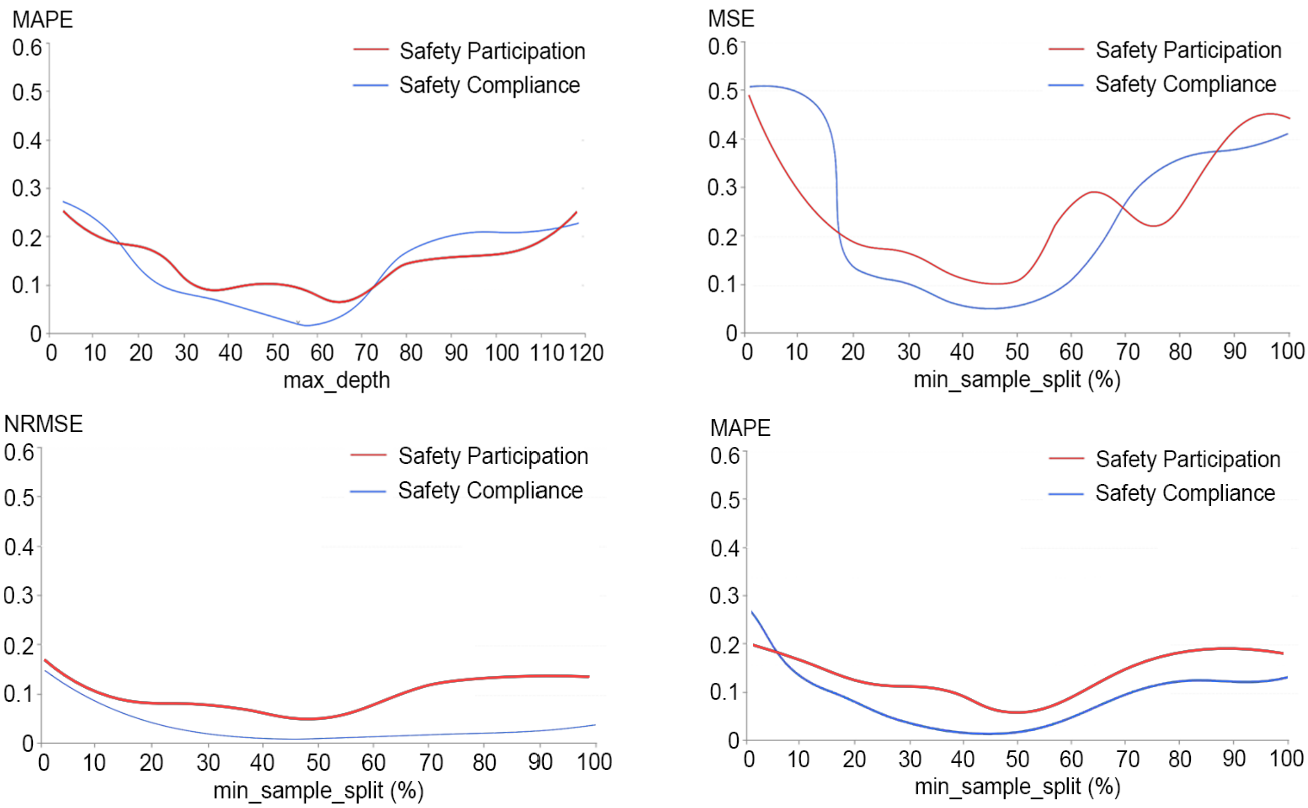

Figure 3. *max_depth*, *min_sample_split*—loss functions (MSE, NRMSE, and MAPE) diagram.

5.2.3. SVM

To fit a set of input–output training data [x_i, y_i] (where x_i = input data, y_i = output data, i = the number of data points), SVM maps the input data's original space into a higher dimensional feature space using a nonlinear Gaussian kernel function $\varphi(x, x') = exp\left(-\sqrt{x^2 - x'^2}/2\sigma^2\right)$ [55]. The function indicates that for any two input instances, x and x', the Euclidean distance (i.e., $\sqrt{(x - x')^2}$) of their corresponding points as mapped in the higher dimensional space is less than the value of σ^2. In the higher dimensional feature space, the task becomes the construction of an optimal surface $y = \omega\varphi(x) + b$ to fit the input–output data, where ω is the weight and b is the bias [55].

To find the best fit, SVM was trained to mathematically identify the optimal σ^2, ω, and b values that indicated the minimal prediction loss on MSE. As suggested in the literature, the parameters σ^2, ω, and b have no defined value range limits [107]. According to the training results presented in Figure 4, it was found that (1) for safety compliance prediction, $\sigma^2 = 4.5$, $\omega = 11.3$, and $b = 4.4$ exhibited the optimal performance among its alternatives on the loss functions (MSE = 0.113, NRMSE = 0.043, and MAPE = 0.047); (2) for safety participation prediction, $\sigma^2 = 6.9$, $\omega = 11.8$, and $b = 0.1$ exhibited the optimal performance among its alternatives on the loss functions (MSE = 0.107, NRMSE = 0.037, MAPE = 0.047). The above values of σ^2, ω, and b were thus used for the SVM model in this research.

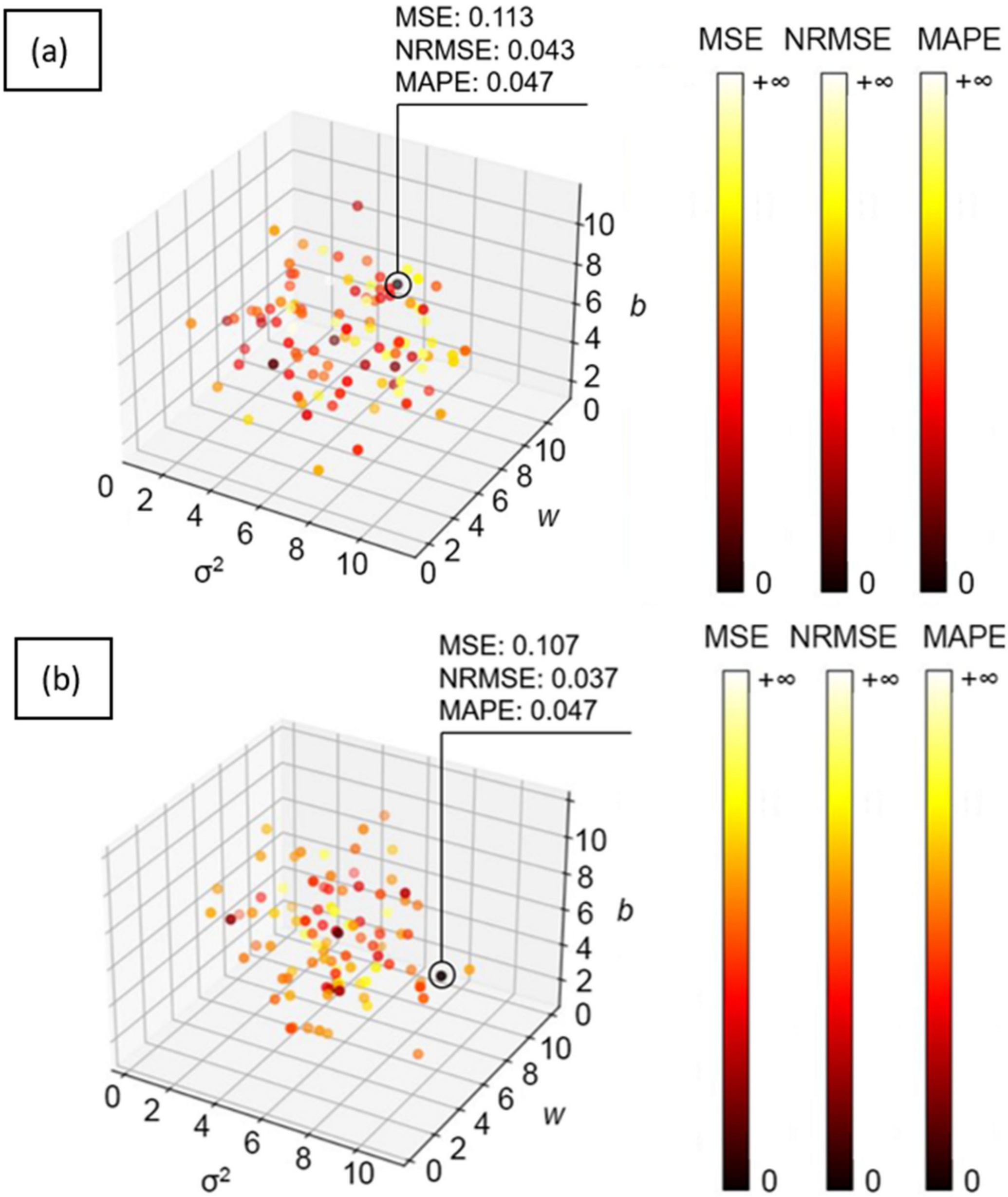

Figure 4. σ^2, ω, b —loss functions (MSE, NRMSE, and MAPE) diagram: (**a**) for safety compliance prediction; (**b**) for safety participation prediction.

5.2.4. KNN

KNN analyses the Euclidean distance between a new observation and all the existing data points in the training set (Figure 5) and assigns the new observation to the respective class that indicates the close points as well as the most similar features [54].

$$Euclidean\ Distance: \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{8}$$

where k = the number of data points to be considered as the close neighbours to a new observation; x_i = feature values of a new observation (i.e., personality scales in this research); and y_i = feature values of an existing data point in the training dataset (i.e., personality scales in this research).

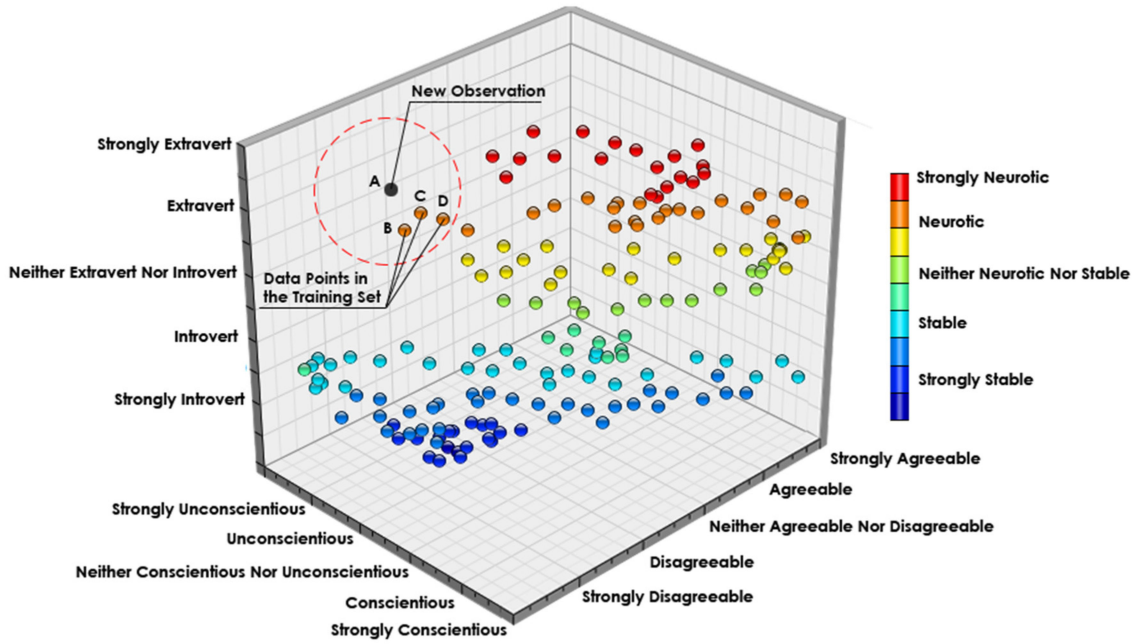

Figure 5. KNN space.

Training of KNN is done to mathematically identify the optimum k value for a given training dataset, which refers to the number of close neighbours to look at when assigning a new observation to any respective class for decision making. As presented in Figure 6, the training results revealed that for a low value of k, KNN overfitted on the safety compliance and participation datasets, which resulted in high errors on the loss functions MSE, NRMSE, and MAPE. For a high value of k, KNN considered an excessive number of data points as close neighbours, which led to poor performance as well. The loss functions reached minima at values of k = 23 for the safety compliance dataset (MSE = 0.067, NRMSE = 0.020, and MAPE = 0.030) and k = 16 for the safety participation dataset (MSE = 0.075, NRMSE = 0.023, and MAPE = 0.027). The optimum k values were therefore determined for the KNN model (i.e., k = 23 for safety compliance, k = 16 for safety participation).

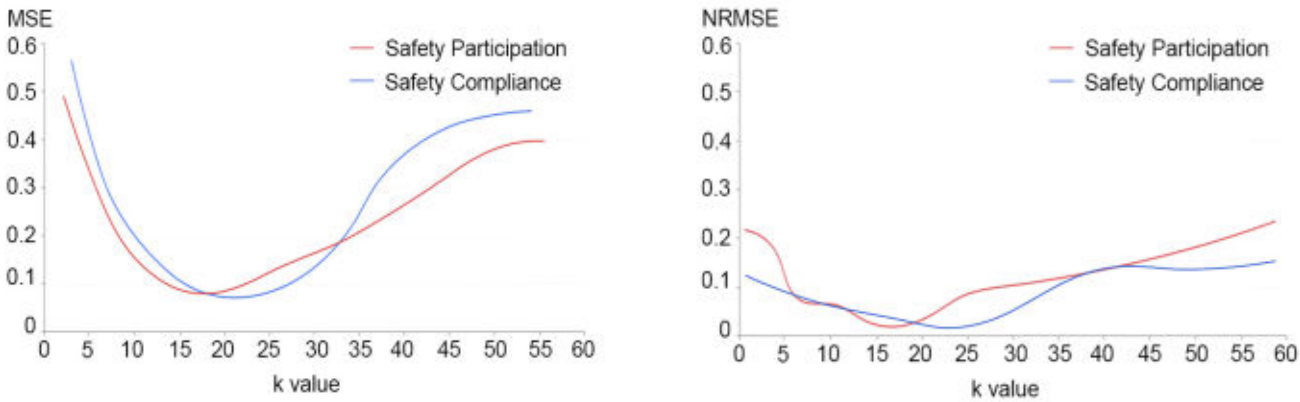

Figure 6. *Cont.*

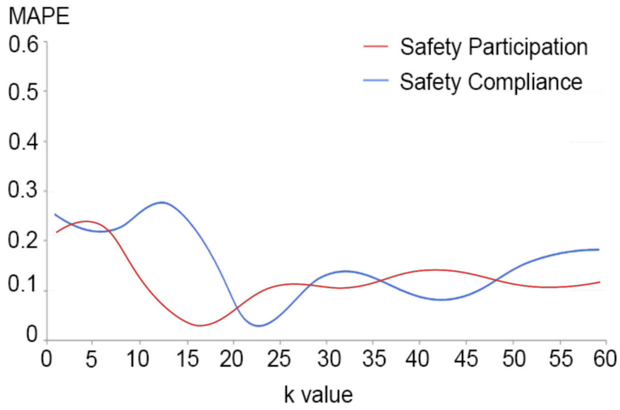

Figure 6. k-value—loss functions (MSE, NRMSE, and MAPE) diagram.

5.2.5. MLR

MLR indicates that the sole output to be predicted is a linear function of one or more explanatory variables [108]: $y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_ix_i$. Gradient descent is an optimisation algorithm that is used to identify the optimal values for the collection of MLR model parameters (b_0, b_1, ... , b_i) [108]. The aim of gradient descent [108] is to identify a point of inflection where $J(b_0,\ b_1,\ \ldots,\ b_i) = \sum(y - \hat{y})^2 = \sum(b_0 + b_1x_1 + b_2x_2 + \cdots + b_ix_i - \hat{y})^2$ yields a minimum (y = MLR prediction given training inputs; $\hat{y}$ = actual value in the training examples). To identify the point of inflection, the gradient of $J(b_0,\ b_1,\ \ldots,\ b_i)$ was determined as follows:

$$\nabla J(b_0,\ b_1,\ \ldots,\ b_i) = \left(\frac{\partial J}{\partial b_0}, \frac{\partial J}{\partial b_1}, \ldots, \frac{\partial J}{\partial b_i}\right) \quad (9)$$

In this research, the architecture of the MLR prediction of construction workers' intended safety behaviour was as follows:

$$y_1 = b_{1,0} + b_{1,1}x_1 + b_{1,2}x_2 + b_{1,3}x_3 + b_{1,4}x_4 \quad (10)$$

$$y_2 = b_{2,0} + b_{2,1}x_1 + b_{2,2}x_2 + b_{2,3}x_3 + b_{2,4}x_4 \quad (11)$$

where y_1 = safety compliance; y_2 = safety participation; x_1 = extraversion; x_2 = agreeableness; x_3 = conscientiousness; x_4 = neuroticism; $b_{1,1}$, $b_{1,2}$, $b_{1,3}$, $b_{1,4}$, $b_{2,1}$, $b_{2,2}$, $b_{2,3}$, $b_{2,4}$ = weights of x_1, x_2, x_3, x_4 in the prediction of y_1 and y_2; $b_{1,0}$, $b_{2,0}$ = constants in the prediction of y_1 and y_2.

Through the process of investigating point of inflection using the training dataset, the minimum of $J(b_0,\ b_1,\ \ldots,\ b_i)$ was identified to be 43.20 and 52.87 for the safety compliance and safety participation datasets, respectively. The corresponding weights and constants were as follows: $b_{1,0} = 0.658$, $b_{1,1} = -0.086$, $b_{1,2} = 0.554$, $b_{1,3} = 0.543$, $b_{1,4} = -0.166$, $b_{2,0} = 1.011$, $b_{2,1} = -0.132$, $b_{2,2} = 0.755$, $b_{2,3} = 0.343$, $b_{2,4} = -0.170$.

5.3. Prediction Performance

After carrying out the optimisation on the ML algorithms, the prediction performance (i.e., prediction loss) of the algorithms was evaluated on the testing dataset. The results for ML predictions on safety compliance and participation intentions are plotted in Figure 7, showing a comparison between the predicted results and actual results (i.e., workers' self-reported intentions to engage in unsafe behaviours) for all the testing samples. The horizontal axis of the plot outlines 40 samples in the testing dataset, and the vertical axis represents one's score on safety compliance and participation intentions, where a higher score indicates better performance in terms of intending to behave safely at work.

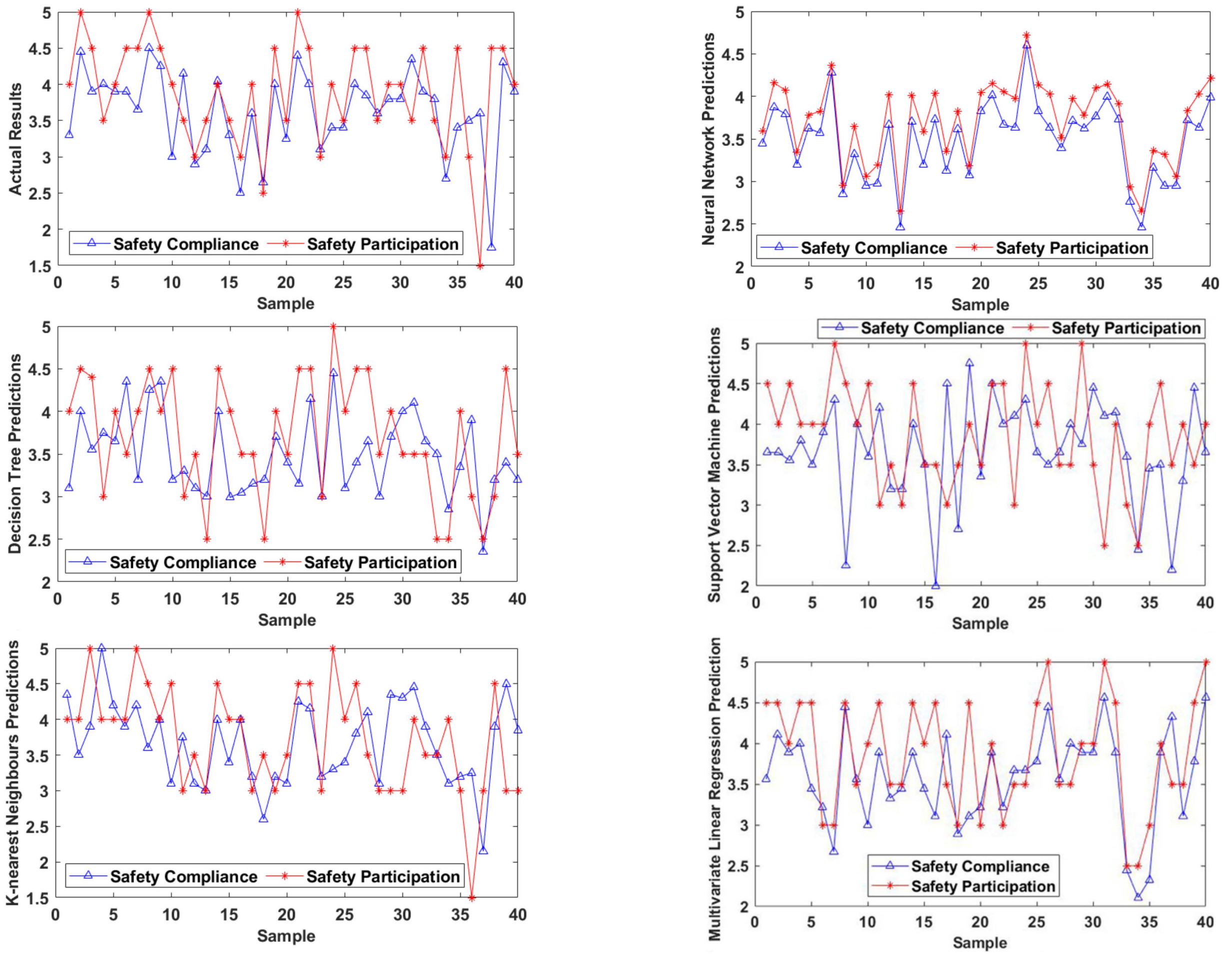

Figure 7. Actual–predicted plot (safety compliance and participation intentions).

Based on the plot (Figure 7), the prediction performance of the ML algorithms was evaluated using the loss functions MSE, NRMSE, and MAPE (1)–(3), which are measures of how accurate a trained algorithm's predictions are compared to the actual results for previously unseen data [101]. The resulting MSE, NRMSE, and MAPE values are presented in Table 7. As can be seen, BP-NN continued the outstanding performance to yield the lowest errors in predicting both safety compliance and participation intentions for the testing samples. According to the threshold, as mentioned (MSE < 0.25, NRMSE < 0.1, and MAPE < 0.1), the prediction performance of BP-NN, DT, SVM, and KNN was found to be highly satisfactory, whereas MLR revealed a less-than-satisfactory performance in the prediction of both safety compliance and participation intention.

Table 7. MSE, NRMSE, and MAPE for predictions on safety compliance and participation intentions.

ML Algorithms	Safety Compliance Intention			Safety Participation Intention		
	MSE	NRMSE	MAPE	MSE	NRMSE	MAPE
NN	0.047	0.017	0.027	0.080	0.023	0.030
DT	0.061	0.021	0.033	0.110	0.030	0.045
SVM	0.107	0.041	0.037	0.093	0.027	0.041
KNN	0.083	0.023	0.042	0.087	0.027	0.032
MLR	0.431	0.159	0.162	0.616	0.174	0.193

Given the five ML algorithms' performance on both the training and testing datasets, it was substantiated that:

1. NN outperformed DT, SVM, KNN, and MLR in capturing the modelling framework of construction workers' personality traits and intentions to engage in safety-related behaviours.
2. NN, DT, SVM, and KNN all showed highly satisfactory performance in predicting both safety compliance and participation intentions.
3. MLR yielded greater prediction errors, which significantly exceeded the upper thresholds (MSE < 0.25, NRMSE < 0.1, and MAPE < 0.1). This implies that a nonlinear relationship exists between construction workers' personality traits and safety behavioural intentions, as the linear method MLR was unable to satisfactorily capture the relationship.

5.4. Weight Analysis

As interpreted in the Introduction section, it is not a single trait but a configuration of traits that constitute individuals' behavioural intentions. Previous studies [1,22] have either been descriptive in nature or provided insufficient insights into the weight of each trait in the personality configuration for workers' safety behavioural intentions, which prevented these studies from establishing a robust predictive basis. This research tested five ML algorithms, with BP-NN producing the best performance in modelling the predictive relationship between workers' personality configuration and behavioural intentions. In order to provide further insights, weight analysis was performed on BP-NN, applying Equation (12), as suggested by Garson [109] and presented below. This method is used to analyse the weight of each input variable in the configuration of a BP-NN for predicting the outputs [109], which has been widely utilised in today's BP-NN research and proven to be highly effective and reliable e.g., [110,111].

$$w_{j,i} = \frac{\sum_{k=1}^{N_k}\left[\left(\frac{\left|w_{j,k}^{IH}\right|}{\sum_{j=1}^{N_j}\left|w_{j,k}^{IH}\right|}\right) \times \left|w_{k,i}^{HO}\right|\right]}{\sum_{j=1}^{N_j}\left\{\sum_{k=1}^{N_k}\left[\left(\frac{\left|w_{j,k}^{IH}\right|}{\sum_{j=1}^{N_j}\left|w_{j,k}^{IH}\right|}\right) \times \left|w_{k,i}^{HO}\right|\right]\right\}} \tag{12}$$

where $w_{j,i}$ = the weight of the jth input variable on the ith output variable; the definitions and values of N_j, N_k, $w_{j,k}^{IH}$, and $w_{k,i}^{HO}$ can be found in Table 8.

Table 8. BP-NN weights.

Input/Output Neurons (I_j/O_i)	Weights	Hidden Neurons (Hk)							
		H_1	H_2	H_3	H_4	H_5	H_6	H_7	H_8
I_1	$w_{1,k}^{IH}$	0.617	0.100	−0.459	0.065	3.925	1.434	3.006	−2.514
I_2	$w_{2,k}^{IH}$	−4.368	1.439	−2.976	0.342	3.888	−0.671	1.456	1.332
I_3	$w_{3,k}^{IH}$	1.888	−1.036	2.561	−0.369	−0.734	−8.384	−1.224	1.200
I_4	$w_{4,k}^{IH}$	2.222	−0.722	1.070	−0.186	−1.412	−0.783	3.461	0.957
O_1	$w_{k,1}^{HO}$	−0.469	2.493	0.325	−5.334	0.200	1.535	0.038	1.301
O_2	$w_{k,2}^{HO}$	−0.717	−0.363	−1.189	−4.906	0.284	4.369	−0.826	3.982

Note: N_j = the number of input neurons (N_j = 4); N_k = the number of hidden neurons (N_k = 8); $w_{j,k}^{IH}$ = weight between the jth input neuron and kth hidden neuron; $w_{k,i}^{HO}$ = weight between the kth hidden neuron and ith output neuron.

The weights of each trait on workers' behavioural intentions were computed based on Equation (12) and the values provided in Table 8, and the results are presented in percentage format in Table 9.

Table 9. Weights of personality configuration on workers' safety behavioural intentions.

Personality Configuration	Weights on Safety Behavioural Intentions (%)	
	Safety Compliance Intention	Safety Participation Intention
Extraversion	11.29	18.48
Agreeableness	32.58	24.85
Conscientiousness	38.19	40.59
Neuroticism	17.94	16.08
Overall	100.00	100.00

It can be observed that conscientiousness with weights of 38.19% and 40.59% exerts the maximum regulatory effects on workers' safety compliance and participation intentions. This result aligns well with the findings of existing meta-analyses of personality traits and occupational safety [27,112]. One possible explanation for this finding may lie in the nature of conscientiousness itself. Conscientiousness is the personality trait of being responsible [71]. The responsibility-oriented nature of conscientiousness leads individuals to consistently be responsible for both their own and team members' well-being and behave safely at work [27]. Such a personality trait would result in the highest involvement in safety behaviour [29,56,113]. It can also be observed that extraversion, with weights of 11.29% and 18.48%, exerts minimal regulatory effects on workers' safety compliance and participation intentions. This result is in agreement with the findings of previous studies in the occupational and traffic safety field [113,114]. According to these studies [113,114], extraversion constitutes a relatively less significant factor in workplace unsafe behaviours compared to the other personality traits (agreeableness, neuroticism, and conscientiousness). Lajunen [113] analysed that a possible reason for this finding is that extraversion is characterised as the extent to which an individual is energetic/active, which implies how susceptible to fatigue they are [27]. Extraversion-related human errors were found to occur when fatigue symptoms start to develop, which are at times compared to other traits that can exert time-continuous influences (e.g., conscientiousness) [113,114]. This is also reflected in industrial statistics. According to the National Safety Council (NSC), fatigue accounted for a small portion of construction injuries (less than 5%) in 2019 in the United States [115].

In addition to conscientiousness and extraversion, the predictive basis also consisted of agreeableness and neuroticism in the configuration, with the weight values being 32.58% and 24.85%, and 17.94%, and 16.08%, respectively (see Table 9). Previous studies have provided insights to explain why agreeableness and neuroticism play significant roles in the prediction of workers' intentions to engage in safety-related behaviours [27,116]. According to the theory of purposeful work behaviour [116], personality traits set individuals out on different intention trajectories in terms of attaining certain goals in the workplace. Agreeableness induces a craving for the goal of communion, where individuals with a high level of agreeableness were found to express a reduced intention to perform unsafely, as unsafe behaviours could put their colleagues' well-being at risk and damage interpersonal relationships as a result [27,116]. Neuroticism reduces the craving for the goal of self-control, where individuals with a high level of neuroticism were found to express a stronger intention to behave impulsively in the workplace, which could result in risky actions [27,116].

6. Discussion

6.1. Theoretical Implications

In the literature, extensive empirical evidence has shown that (1) the formation of human intentions has personality origins [27,35] and (2) individual variation in trait scores is reflected upon different scales measuring individuals' intentions about behaving safely (e.g., "strongly proactive", "proactive", "neither proactive nor negative", "reactive", "strongly reactive") [25,49–52]. Based on the findings, it has been suggested that the intention to engage in safety-related behaviours at the individual level can be predicted on the basis of quantifying individual variation in personality [16,47]. However, a dearth of research in the construction sector provides little empirical evidence. The studies by [22] and Zhang, Xiang, Zhang, Chen and Ren [1] called attention to this field. However, their studies are descriptive in nature and provide insufficient insights into the constitution of personality predictors for workers' behavioural intentions, which prevents these studies from reaching an operative prediction framework. This research is an extension of the existing studies, which addresses the knowledge gap by providing a ML-based approach to help understand workers' intentions to engage in unsafe behaviours by quantifying workers' individual variations in personality.

There are two other knowledge gaps that are addressed in this research. First, this research finds that a nonlinear relationship exists between construction workers' personality traits and inner intentions to engage in unsafe behaviours, as the linear method MLR was unable to satisfactorily capture the relationship and yielded greater prediction errors than the other nonlinear ML methods. Second, this research gains an enhanced understanding of the weights of each personality trait in the constitution of prediction practice for workers' intentions to engage in unsafe behaviours.

In addition, the authors have reviewed papers that aim at predicting human behaviour in other areas such as criminal behaviour using ML techniques [113,117,118]. Our model did not have higher accuracy compared to these studies, as they have all shown fair prediction outcomes. However, our research has a technical advantage in the method design compared to these studies. These studies [113,117,118], only utilised one ML technique to develop the prediction model (e.g., k-means clustering). In our study, five ML architectures—backpropagation neural network, decision tree, support vector machine, k-nearest neighbours, and multivariate linear regression—were used to capture the predictive relationship between construction workers' personalities and their intentions to engage in unsafe behaviours, and the five models were compared. Backpropagation neural network outperformed other algorithms, yielding minimal prediction loss, and was determined to be the best approach.

6.2. Practical Implications

The BP-NN approach can be utilised to generate quantifiable predictions to help understand the extent of workers' intentions to engage in unsafe behaviours. According to the decisions of the ML approach, a "low safety compliance" prediction implies a stronger intention of engaging in non-compliant work behaviours such as taking shortcuts, not using safety equipment, and violating safety procedures and rules. A "low safety participation" prediction implies reduced intent to perform solicitous safety activities such as reporting co-workers' safety problems, keeping the workplace clean, and caring for colleagues' safety. Such knowledge is useful for understanding undesirable aspects in different workers (at the individual level) in order to recommend suitable preventive strategies for workers with different needs. Preventive strategies can be safety training programmes to promote individuals' behavioural adherence to safety standards and desire to contribute to organisational safety. For individuals with different "safety compliance" and "safety participation" scores (predicted), variation in safety training in terms of the range of intensity can be implemented. Therefore, this study highlights the significance of predicting workers' inner intentions to engage in unsafe behaviours for the design of safety training programmes. In the next step of this research, the safety training intensity taxonomy in relation to individuals' prediction scores on "safety compliance" and "safety participation" will be further studied.

7. Conclusions

This research presents a machine learning (ML)-based approach to help understand workers' inner intentions to engage in unsafe behaviours by quantifying workers' individual variations in personality. The ML approach is based on the backpropagation neural network (BP-NN) architecture and utilises the fundamental configuration of human personality (i.e., extraversion, neuroticism, agreeableness, and conscientiousness) to predict workers' intentions, which consist of safety compliance and safety participation dimensions.

The ML approach performed satisfactorily on the cross-validation test, yielding minimal prediction loss in the loss functions MSE, NRMSE, and MAPE for both the training and testing datasets (training dataset: MSE = 0.042, NRMSE: 0.010, and MAPE: 0.021; testing dataset: MSE = 0.047, NRMSE = 0.017, and MAPE = 0.030). Weight analysis was then performed on BP-NN to gain insights into the weights of each trait in the personality configuration for workers' intentions to engage in safety-related behaviours, which constitutes the foundation for the modelling and prediction framework. The results showed that the weights of conscientiousness, agreeableness, neuroticism, and extraversion on workers' safety compliance and participation intentions are 38.19% and 40.59%, 32.58% and 24.85%, 17.94% and 16.08%, and 11.29% and 18.49%, respectively. The ML approach can be utilised to generate quantifiable predictions to help understand the extent of workers' intentions to engage in unsafe behaviours. Such knowledge is useful for understanding undesirable aspects in different workers (at the individual level) in order to recommend suitable preventive strategies for workers with different needs.

There are limitations to this research. First, due to resource constraints only a limited number of construction projects in NZ were visited to collect data. The sampling process might leave out sections of the population that could be significant to this study. Nevertheless, the researchers attempted to overcome this issue by approaching workers from as many construction sites as possible. It would be interesting for future research to expand the data collection from both New Zealand and the international construction community to test the ML approach. Second, as mentioned previously, for individuals with different "safety compliance" and "safety participation" scores (predicted), variation in safety training in terms of the range of intensity can be implemented. In the next step of this research, the safety training intensity taxonomy in relation to individuals' prediction scores on "safety compliance" and "safety participation" will be further studied.

Author Contributions: Conceptualization, Y.G. and V.A.G.; methodology, Y.G. and V.A.G.; software, Y.G.; validation, Y.G.; formal analysis, Y.G.; investigation, Y.G.; resources, Y.G. and V.A.G.; data curation, Y.G.; writing—original draft preparation, Y.G.; writing—review and editing, Y.G., V.A.G., T.W.Y., G.C.-G. and R.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE (protocol code 019515 and date of approval 16 June 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ML	Machine Learning
BP-NN	Backpropagation Neural Network
DT	Decision Tree
SVM	Support Vector Machine
KNN	K-nearest Neighbours
MLR	Multivariate Linear Regression
TPB	Theory of Planned Behaviour
SDT	Self-determination Theory
PPE	Personal Protective Equipment
CNN	Convolutional Neural Network

MSE	Mean Squared Error
NRMSE	Normalized Root Mean Square Error
MAPE	Mean Absolute Percentage Error
BFI	Big Five Inventory
NEO-FFI	NEO Five-Factor Inventory
NEO-PI-R	Revised NEO Personality Inventory
FOR-BFI	Frame-of-Reference Big Five Inventory
SBS	Safety Behaviour Scale
SDQ	Safety Diagnosis Questionnaire
SCI-WHSB	Safety Climate Index: Workers' Health and Safety Behaviour
WSS	Work Safety Scale
SPS	Safety Performance Scale
NSC	National Safety Council

References

1. Zhang, J.; Xiang, P.; Zhang, R.; Chen, D.; Ren, Y. Mediating Effect of Risk Propensity between Personality Traits and Unsafe Behavioral Intention of Construction Workers. *J. Constr. Eng. Manag.* **2020**, *146*, 04020023. [CrossRef]
2. BLS. U.S. Bureau of Labor Statistics: Industry Injuries, Illnesses, and Fatalities Statistics. Available online: https://www.bls.gov/iif/oshover.htm (accessed on 16 May 2022).
3. HSE. UK Health and Safety Executive: Statistics on Nonfatal and Fatal Injuries in the Workplace in Great Britain. Available online: http://www.hse.gov.uk/statistics/tables/index.htm (accessed on 16 May 2022).
4. Gao, Y.; Gonzalez, V.A.; Yiu, T.W. The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: A systematic review. *Comput. Educ.* **2019**, *138*, 101–115. [CrossRef]
5. Li, X.; Li, H.; Skitmore, M.; Wang, F. Understanding the influence of safety climate and productivity pressure on non-helmet use behavior at construction sites: A case study. *Eng. Constr. Archit. Manag.* **2021**. [CrossRef]
6. Gondia, A.; Siam, A.; El-Dakhakhni, W.; Nassar, A.H. Machine Learning Algorithms for Construction Projects Delay Risk Prediction. *J. Constr. Eng. Manag.* **2020**, *146*, 04019085. [CrossRef]
7. Leu, S.-S.; Chang, C.-M. Bayesian-network-based safety risk assessment for steel construction projects. *Accid. Anal. Prev.* **2013**, *54*, 122–133. [CrossRef] [PubMed]
8. Patel, D.A.; Jha, K.N. Neural Network Model for the Prediction of Safe Work Behavior in Construction Projects. *J. Constr. Eng. Manag.* **2015**, *141*, 04014066. [CrossRef]
9. Jitwasinkul, B.; Hadikusumo, B.H.W.; Memon, A.Q. A Bayesian Belief Network model of organizational factors for improving safe work behaviors in Thai construction industry. *Saf. Sci.* **2016**, *82*, 264–273. [CrossRef]
10. Patel, D.A.; Jha, K.N. Evaluation of construction projects based on the safe work behavior of co-employees through a neural network model. *Saf. Sci.* **2016**, *89*, 240–248. [CrossRef]
11. Patel, D.A.; Jha, K.N. Neural Network Approach for Safety Climate Prediction. *J. Manag. Eng.* **2015**, *31*, 05014027. [CrossRef]
12. Yule, S.; Flin, R.; Murdy, A. The role of management and safety climate in preventing risk-taking at work. *Int. J. Risk Assess. Manag.* **2007**, *7*, 137–151. [CrossRef]
13. Earle, A.M.; Napper, L.E.; LaBrie, J.W.; Brooks-Russell, A.; Smith, D.J.; de Rutte, J. Examining interactions within the theory of planned behavior in the prediction of intentions to engage in cannabis-related driving behaviors. *J. Am. Coll. Health* **2020**, *68*, 374–380. [CrossRef]
14. Moiron, M.; Laskowski, K.L.; Niemelä, P.T. Individual differences in behaviour explain variation in survival: A meta-analysis. *Ecol. Lett.* **2020**, *23*, 399–408. [CrossRef]
15. Guerin, R.J.; Toland, M.D. An application of a modified theory of planned behavior model to investigate adolescents' job safety knowledge, norms, attitude and intention to enact workplace safety and health skills. *J. Saf. Res.* **2020**, *72*, 189–198. [CrossRef]
16. Ajzen, I. From Intentions to Actions: A Theory of Planned Behavior. In *Action Control: From Cognition to Behavior*; Kuhl, J., Beckmann, J., Eds.; Springer: Berlin/Heidelberg, Germany, 1985; pp. 11–39. [CrossRef]
17. Ryan, R.M.; Deci, E.L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **2000**, *55*, 68–78. [CrossRef]
18. Ahmed, E.; Ward, R. Analysis of factors influencing acceptance of personal, academic and professional development e-portfolios. *Comput. Hum. Behav.* **2016**, *63*, 152–161. [CrossRef]
19. Blumstein, D.T. Applying behavioral ecology: A comment on Harding et al. *Behav. Ecol.* **2019**, *30*, 1512. [CrossRef]
20. Mahdavi, A. Explanatory stories of human perception and behavior in buildings. *Build. Environ.* **2020**, *168*, 106498. [CrossRef]
21. Thys, B.; Eens, M.; Aerts, S.; Delory, A.; Iserbyt, A.; Pinxten, R. Exploration and sociability in a highly gregarious bird are repeatable across seasons and in the long term but are unrelated. *Anim. Behav.* **2017**, *123*, 339–348. [CrossRef]
22. Wong, D.B.; Lee, S.G. Modelling the predictors of intention in workplace safety compliance of a multi-ethnic workforce. *Saf. Sci.* **2016**, *88*, 155–165. [CrossRef]
23. Abdelhamid, T.S.; Everett, J.G. Identifying Root Causes of Construction Accidents. *J. Constr. Eng. Manag.* **2000**, *126*, 52–60. [CrossRef]

24. Feng, Z.; González, V.A.; Amor, R.; Lovreglio, R.; Cabrera-Guerrero, G. Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Comput. Educ.* **2018**, *127*, 252–266. [CrossRef]
25. Laurent, J.; Chmiel, N.; Hansez, I. Personality and safety citizenship: The role of safety motivation and safety knowledge. *Heliyon* **2020**, *6*, e03201. [CrossRef]
26. Tement, S.; Plohl, N.; Horvat, M.; Musil, B.; Jakus, G.; Sodnik, J.J.A. Is past behavior the best predictor of future behavior? Comparing the behavioral validity of self-reported speeding and driving simulator data. *Age* **2019**, *30*, 08.
27. Hasanzadeh, S.; Dao, B.; Esmaeili, B.; Dodd Michael, D. Role of Personality in Construction Safety: Investigating the Relationships between Personality, Attentional Failure, and Hazard Identification under Fall-Hazard Conditions. *J. Constr. Eng. Manag.* **2019**, *145*, 04019052. [CrossRef]
28. Gillison, F.B.; Rouse, P.; Standage, M.; Sebire, S.J.; Ryan, R.M. A meta-analysis of techniques to promote motivation for health behaviour change from a self-determination theory perspective. *Health Psychol. Rev.* **2019**, *13*, 110–130. [CrossRef]
29. Beus, J.M.; Dhanani, L.Y.; McCord, M.A. A meta-analysis of personality and workplace safety: Addressing unanswered questions. *J. Appl. Psychol.* **2015**, *100*, 481–498. [CrossRef]
30. Mentrikoski, J.M.; Duncan, C.L.; Enlow, P.T.; Aballay, A.M. Predicting Adolescents' Intentions to Engage in Fire Risk Behaviors: An Application of the Theory of Planned Behavior. *Burns* **2019**, *45*, 1242–1250. [CrossRef] [PubMed]
31. Neves, H.C.C.; Silva e Souza, A.C.; Medeiros, M.; Munari, D.B.; Ribeiro, L.C.M.; Tipple, A.F.V. Safety of nursing staff and determinants of adherence to personal protective equipment. *Rev. Lat.-Am. Enferm.* **2011**, *19*, 354–361. [CrossRef]
32. Rose, A.; Rae, W.I.D. Personal Protective Equipment Availability and Utilization among Interventionalists. *Saf. Health Work* **2019**, *10*, 166–171. [CrossRef] [PubMed]
33. Arsenie, P.; Hanzu-Pazara, R. Human errors and oil pollution from tankers. *Int. J. Mar. Navig. Saf. Sea Transp.* **2008**, *2*, 409–413. Available online: https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-0d29bbbf-d521-4614-b65b-ff88524dc731 (accessed on 16 May 2022).
34. Baldissone, G.; Comberti, L.; Bosca, S.; Murè, S. The analysis and management of unsafe acts and unsafe conditions. Data collection and analysis. *Saf. Sci.* **2019**, *119*, 240–251. [CrossRef]
35. Jiang, Z.; Fang, D.; Zhang, M. Understanding the Causation of Construction Workers' Unsafe Behaviors Based on System Dynamics Modeling. *J. Manag. Eng.* **2015**, *31*, 04014099. [CrossRef]
36. Shakerian, M.; Jahangiri, M.; Alimohammadlou, M.; Nami, M.; Choobineh, A. Individual cognitive factors affecting unsafe acts among Iranian industrial workers: An integrative meta-synthesis interpretive structural modeling (ISM) approach. *Saf. Sci.* **2019**, *120*, 89–98. [CrossRef]
37. Norman, P.; Webb, T.L.; Millings, A. Using the theory of planned behaviour and implementation intentions to reduce binge drinking in new university students. *Psychol. Health* **2019**, *34*, 478–496. [CrossRef]
38. Li, A.S.W.; Figg, G.; Schüz, B. Socioeconomic Status and the Prediction of Health Promoting Dietary Behaviours: A Systematic Review and Meta-Analysis Based on the Theory of Planned Behaviour. *Appl. Psychol. Health Well-Being* **2019**, *11*, 382–406. [CrossRef]
39. Hakim, J.D.; Chami, J.; Keay, K.A. μ-Opioid and dopamine-D2 receptor expression in the nucleus accumbens of male Sprague-Dawley rats whose sucrose consumption, but not preference, decreases after nerve injury. *Behav. Brain Res.* **2020**, *381*, 112416. [CrossRef]
40. Corr, P.J.; Matthews, G. *The Cambridge Handbook of Personality Psychology*; Cambridge University Press: Cambridge, UK, 2009. Available online: https://www.cambridge.org/core/books/cambridge-handbook-of-personality-psychology/8E0820BD8B627FB17B009A34E81A5DEB (accessed on 16 May 2022).
41. Delvecchio, G.; Bellani, M.; Altamura, A.C.; Brambilla, P. The association between the serotonin and dopamine neurotransmitters and personality traits. *Epidemiol. Psychiatr. Sci.* **2016**, *25*, 109–112. [CrossRef]
42. Hennig, J.; Netter, P.; Munk, A.J.L. Interaction between serotonin and dopamine and impulsivity: A gene × gene—Interaction approach. *Personal. Individ. Differ.* **2020**, *169*, 110014. [CrossRef]
43. Fong, M.; Zhao, K.; Smillie, L.D. Personality and competitiveness: Extraversion, agreeableness, and their aspects, predict self-reported competitiveness and competitive bidding in experimental auctions. *Personal. Individ. Differ.* **2020**, *169*, 109907. [CrossRef]
44. Huang, M.-W.; Chiang, T.-A.; Lo, P.-Y.; Huang, C.-S. Relationship among methadone dose, polymorphisms of dopamine D2 receptor and tri-dimensional personality questionnaire in heroin-addicted patients. *Behav. Brain Funct.* **2016**, *12*, 24. [CrossRef]
45. London, E.D. *Human Brain Imaging Links Dopaminergic Systems to Impulsivity*; Springer: Berlin/Heidelberg, Germany, 2020.
46. Hosken, D.J.; Hunt, J.; Wedell, N. *Genes and Behaviour: Beyond Nature-Nurture*; John Wiley & Sons: Hoboken, NJ, USA, 2019.
47. Bell, A.M.; Hankison, S.J.; Laskowski, K.L. The repeatability of behaviour: A meta-analysis. *Anim. Behav.* **2009**, *77*, 771–783. [CrossRef] [PubMed]
48. Barańczuk, U. The five factor model of personality and emotion regulation: A meta-analysis. *Personal. Individ. Differ.* **2019**, *139*, 217–227. [CrossRef]
49. Pletzer, J.L.; Bentvelzen, M.; Oostrom, J.K.; de Vries, R.E. A meta-analysis of the relations between personality and workplace deviance: Big Five versus HEXACO. *J. Vocat. Behav.* **2019**, *112*, 369–383. [CrossRef]

50. Pei, W.; Xiao, D. Curriculum Development and Research on Construction Safety Psychology and Behavior. In Proceedings of the 2017 2nd International Seminar on Education Innovation and Economic Management (SEIEM 2017), Penang, Malaysia, 28–30 December 2017; pp. 270–273. [CrossRef]
51. Fischer, R.; Lee, A.; Verzijden, M.N. Dopamine genes are linked to Extraversion and Neuroticism personality traits. *Sci. Rep.* **2018**, *8*, 1733. [CrossRef]
52. Ray, S.; Fritzon, K. Overview of the Impact of Psychopathy and Other Problematic Personality Constructs in the Workplace. In *Corporate Psychopathy: Investigating Destructive Personalities in the Workplace*; Fritzon, K., Brooks, N., Croom, S., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 199–253. [CrossRef]
53. Isler, L.; Fletcher, G.J.O.; Liu, J.H.; Sibley, C.G. Validation of the four-profile configuration of personality types within the Five-Factor Model. *Personal. Individ. Differ.* **2017**, *106*, 257–262. [CrossRef]
54. Castanier, C.; Scanff, C.L.; Woodman, T. Who Takes Risks in High-Risk Sports? A Typological Personality Approach. *Res. Q. Exerc. Sport* **2013**, *81*, 478–484. [CrossRef]
55. Landay, K.; Wood, D.; Harms, P.D.; Ferrell, B.; Nambisan, S. Relationships between personality facets and accident involvement among truck drivers. *J. Res. Personal.* **2020**, *84*, 103889. [CrossRef]
56. Gao, Y.; Gonzalez, V.A.; Yiu, T.W. Exploring the Relationship between Construction Workers' Personality Traits and Safety Behaviour. *J. Constr. Eng. Manag.* **2020**, *146*, 04019111. [CrossRef]
57. Rau, P.-L.P.; Liao, P.-C.; Guo, Z.; Zheng, J.; Jing, B. Personality factors and safety attitudes predict safety behaviour and accidents in elevator workers. *Int. J. Occup. Saf. Ergon.* **2018**, *1*, 719–727. [CrossRef]
58. Panuwatwanich, K.; Al-Haadir, S.; Stewart, R.A. Influence of safety motivation and climate on safety behaviour and outcomes: Evidence from the Saudi Arabian construction industry. *Int. J. Occup. Saf. Ergon.* **2017**, *23*, 60–75. [CrossRef]
59. Alloghani, M.; Al-Jumeily, D.; Mustafina, J.; Hussain, A.; Aljaaf, A.J. A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. In *Supervised and Unsupervised Learning for Data Science*; Berry, M.W., Mohamed, A., Yap, B.W., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 3–21. [CrossRef]
60. Antwi-Afari, M.F.; Li, H.; Yu, Y.; Kong, L. Wearable insole pressure system for automated detection and classification of awkward working postures in construction workers. *Autom. Constr.* **2018**, *96*, 433–441. [CrossRef]
61. Tümer, A.E.; Edebali, S. Modeling of Trivalent Chromium Sorption onto Commercial Resins by Artificial Neural Network. *Appl. Artif. Intell.* **2019**, *33*, 349–360. [CrossRef]
62. Jorgensen, B. *Theory of Linear Models*; Routledge: New York, NY, USA, 2019. [CrossRef]
63. Linnainmaa, S. The Representation of the Cumulative Rounding Error of an Algorithm as a Taylor Expansion of the Local Rounding Errors. Master's Thesis, University of Helsinki, Helsinki, Finland, 1970.
64. Quinlan, J.R. Induction of decision trees. *Mach. Learn.* **1986**, *1*, 81–106. [CrossRef]
65. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, PA, USA, 27–29 July 1992; pp. 144–152. [CrossRef]
66. Altman, N.S. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *Am. Stat.* **1992**, *46*, 175–185. [CrossRef]
67. Mardia, K.; Kent, J.; Bibby, J. *Multivariate Analysis*; Academic Press Inc.: Cambridge, MA, USA, 1979; Volume 15, p. 518. Available online: https://www.elsevier.com/books/multivariate-analysis/mardia/978-0-08-057047-1 (accessed on 16 May 2022).
68. Sony, S.; Dunphy, K.; Sadhu, A.; Capretz, M. A systematic review of convolutional neural network-based structural condition assessment techniques. *Eng. Struct.* **2021**, *226*, 111347. [CrossRef]
69. Gao, C.; Elzarka, H. The use of decision tree based predictive models for improving the culvert inspection process. *Adv. Eng. Inform.* **2021**, *47*, 101203. [CrossRef]
70. Amin, S. Backpropagation—Artificial Neural Network (BP-ANN): Understanding gender characteristics of older driver accidents in West Midlands of United Kingdom. *Saf. Sci.* **2020**, *122*, 104539. [CrossRef]
71. Hongtao, N. Smart safety early warning model of landslide geological hazard based on BP neural network. *Saf. Sci.* **2020**, *123*, 104572. [CrossRef]
72. Christodoulou, E.; Ma, J.; Collins, G.S.; Steyerberg, E.W.; Verbakel, J.Y.; Van Calster, B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J. Clin. Epidemiol.* **2019**, *110*, 12–22. [CrossRef]
73. Bai, Y.; Chen, Z.; Xie, J.; Li, C. Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models. *J. Hydrol.* **2016**, *532*, 193–206. [CrossRef]
74. Costa, P.T.; McCrae, R.R.; Löckenhoff, C.E. Personality Across the Life Span. *Annu. Rev. Psychol.* **2019**, *70*, 423–448. [CrossRef]
75. Hofmann, D.A.; Burke, M.J.; Zohar, D. 100 years of occupational safety research: From basic protections and work analysis to a multilevel view of workplace safety and risk. *J. Appl. Psychol.* **2017**, *102*, 375–388. [CrossRef]
76. John, O.P.; Donahue, E.M.; Kentle, R.L. The Big Five Inventory—Versions 4a and 54. Available online: https://www.ocf.berkeley.edu/~{}johnlab/bfi.php (accessed on 16 May 2022).
77. Costa, P.T.; McCrae, R.R. NEO Five-Factor Inventory (NEO-FFI). Available online: https://www.parinc.com/Products/Pkey/274 (accessed on 16 May 2022).
78. Costa, P.T.; McCrae, R.R. *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional Manual*; Psychological Assessment Resources, Incorporated: Lutz, FL, USA, 1992. Available online: https://books.google.co.nz/books?id=D5CdtwAACAAJ (accessed on 16 May 2022).

79. Schmit, M.J.; Ryan, A.M.; Stierwalt, S.L.; Powell, A.B. Frame-of-Reference Effects on Personality Scale Scores and Criterion-Related Validity. *J. Appl. Psychol.* **1995**, *80*, 607–620. [CrossRef]
80. Soutter, A.R.B.; Bates, T.C.; Mõttus, R. Big Five and HEXACO Personality Traits, Proenvironmental Attitudes, and Behaviors: A Meta-Analysis. *Perspect. Psychol. Sci.* **2020**, *15*, 913–941. [CrossRef]
81. Feher, A.; Vernon, P.A. Looking beyond the Big Five: A selective review of alternatives to the Big Five model of personality. *Personal. Individ. Differ.* **2020**, *169*, 110002. [CrossRef]
82. Kerry, N.; Murray, D.R. Strong personalities: Investigating the relationships between grip strength, self-perceived formidability, and Big Five personality traits. *Personal. Individ. Differ.* **2018**, *131*, 216–221. [CrossRef]
83. Hamby, T.; Taylor, W.; Snowden, A.K.; Peterson, R.A. A Meta-Analysis of the Reliability of Free and For-Pay Big Five Scales. *J. Psychol.* **2016**, *150*, 422–430. [CrossRef]
84. Kunnel John, R.; Xavier, B.; Waldmeier, A.; Meyer, A.; Gaab, J. Psychometric Evaluation of the BFI-10 and the NEO-FFI-3 in Indian Adolescents. *Front. Psychol.* **2019**, *10*, 1057. [CrossRef] [PubMed]
85. Beckmann, N.; Birney, D.P.; Beckmann, J.F.; Wood, R.E.; Sojo, V.; Bowman, D. Inter-individual differences in intra-individual variability in personality within and across contexts. *J. Res. Personal.* **2020**, *85*, 103909. [CrossRef]
86. Anglim, J.; O'Connor, P. Measurement and research using the Big Five, HEXACO, and narrow traits: A primer for researchers and practitioners. *Aust. J. Psychol.* **2019**, *71*, 16–25. [CrossRef]
87. Swift, V.; Peterson, J.B. Contextualization as a means to improve the predictive validity of personality models. *Personal. Individ. Differ.* **2019**, *144*, 153–163. [CrossRef]
88. Simonet, D.V.; Castille, C.M. The search for meaningful work: A network analysis of personality and the job characteristics model. *Personal. Individ. Differ.* **2020**, *152*, 109569. [CrossRef]
89. Woo, S.E.; Jin, J.; LeBreton, J.M. Specificity Matters: Criterion-Related Validity of Contextualized and Facet Measures of Conscientiousness in Predicting College Student Performance. *J. Personal. Assess.* **2015**, *97*, 301–309. [CrossRef]
90. Hirschfeld, R.R.; Lawson, L. Predicting Individuals' Interest After a Performance Result: The Roles of Motivational Orientations at High and Low Performance. *J. Appl. Soc. Psychol.* **2008**, *38*, 557–584. [CrossRef]
91. Wang, Q.; Bowling, N.A. A Comparison of General and Work-specific Personality Measures as Predictors of Organizational Citizenship Behavior. *Int. J. Sel. Assess.* **2016**, *24*, 172–188. [CrossRef]
92. Shaffer, J.A.; Postlethwaite, B.E. A Matter of Context: A Meta-Analytic Investigation of the Relative Validity of Contextualized and Noncontextualized Personality Measures. *Pers. Psychol.* **2012**, *65*, 445–494. [CrossRef]
93. Smith, R.W.; DeNunzio, M.M. Examining personality—Job characteristic interactions in explaining work outcomes. *J. Res. Personal.* **2020**, *84*, 103884. [CrossRef]
94. Hayes, B.E.; Perander, J.; Smecko, T.; Trask, J. Measuring perceptions of workplace safety: Development and validation of the work safety scale. *J. Saf. Res.* **1998**, *29*, 145–161. [CrossRef]
95. Hoyos, C.G.; Ruppert, F. Safety diagnosis in industrial work settings: The safety diagnosis questionnaire. *J. Saf. Res.* **1995**, *26*, 107–117. [CrossRef]
96. Garcia, A.M.; Boix, P.; Canosa, C. Why do workers behave unsafely at work? Determinants of safe work practices in industrial workers. *Occup. Environ. Med.* **2004**, *61*, 239–246. [CrossRef]
97. Mearns, K.; Flin, R.; Gordon, R.; Fleming, M. Human and organizational factors in offshore safety. *Work Stress* **2001**, *15*, 144–160. [CrossRef]
98. Griffin, M.A.; Neal, A. Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *J. Occup. Health Psychol.* **2000**, *5*, 347–358. [CrossRef] [PubMed]
99. Singh, A.; Misra, S.C. A Dominance based Rough Set analysis for investigating employee perception of safety at workplace and safety compliance. *Saf. Sci.* **2020**, *127*, 104702. [CrossRef]
100. Müller, S.L.; Shehadeh, M.A.; Schröder, S.; Richert, A.; Jeschke, S. An overview of work analysis instruments for hybrid production workplaces. *AI Soc.* **2018**, *33*, 425–432. [CrossRef]
101. Ugwu, F.O.; Idike, A.N.; Ibiam, O.E.; Akwara, F.A.; Okorie, C.O. Transformational leadership and management safety practices: Their role in the relationship between work pressure and compliance with safety work behaviour in a health-care sector industry. *J. Psychol. Afr.* **2020**, *30*, 1–8. [CrossRef]
102. Stratman, J.L.; Youssef-Morgan, C.M. Can positivity promote safety? Psychological capital development combats cynicism and unsafe behavior. *Saf. Sci.* **2019**, *116*, 13–25. [CrossRef]
103. Wen Lim, H.; Li, N.; Fang, D.; Wu, C. Impact of Safety Climate on Types of Safety Motivation and Performance: Multigroup Invariance Analysis. *J. Manag. Eng.* **2018**, *34*, 04018002. [CrossRef]
104. Reddy, A.S.B.; Juliet, D.S. Transfer Learning with ResNet-50 for Malaria Cell-Image Classification. In Proceedings of the 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 4–6 April 2019; pp. 0945–0949. [CrossRef]
105. Krzykowska, K.; Krzykowski, M. Forecasting Parameters of Satellite Navigation Signal through Artificial Neural Networks for the Purpose of Civil Aviation. *Int. J. Aerosp. Eng.* **2019**, *2019*, 7632958. [CrossRef]
106. Singaravel, S.; Suykens, J.; Geyer, P. Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction. *Adv. Eng. Inform.* **2018**, *38*, 81–90. [CrossRef]

107. Desai, S.D.; Giraddi, S.; Narayankar, P.; Pudakalakatti, N.R.; Sulegaon, S. Back-Propagation Neural Network Versus Logistic Regression in Heart Disease Classification. In *Advanced Computing and Communication Technologies*; Springer: Singapore, 2019; pp. 133–144.
108. Lafif Tej, M.; Holban, S. Determining Optimal Multi-layer Perceptron Structure Using Linear Regression. In *Business Information Systems—BIS 2019*; Springer: Cham, Switzerland, 2019; pp. 232–246.
109. Velasco, L.C.P.; Estoperez, N.R.; Jayson, R.J.R.; Sabijon, C.J.T.; Sayles, V.C. Performance Analysis of Artificial Neural Networks Training Algorithms and Activation Functions in Day-Ahead Base, Intermediate, and Peak Load Forecasting. In *Advances in Information and Communication*; Springer: Cham, Switzerland, 2020; pp. 284–298. [CrossRef]
110. Scardapane, S.; Van Vaerenbergh, S.; Totaro, S.; Uncini, A. Kafnets: Kernel-based non-parametric activation functions for neural networks. *Neural Netw.* **2019**, *110*, 19–32. [CrossRef]
111. Yang, L.; Liu, S.; Tsoka, S.; Papageorgiou, L.G. A regression tree approach using mathematical programming. *Expert Syst. Appl.* **2017**, *78*, 347–357. [CrossRef]
112. Liu, M.; Zou, L.; Yu, X.; Zhou, Y.; Wang, X.; Tang, B. Knowledge Aided Covariance Matrix Estimation via Gaussian Kernel Function for Airborne SR-STAP. *IEEE Access* **2020**, *8*, 5970–5978. [CrossRef]
113. Lajunen, T. Personality and accident liability: Are extraversion, neuroticism and psychoticism related to traffic and occupational fatalities? *Personal. Individ. Differ.* **2001**, *31*, 1365–1373. [CrossRef]
114. Garson, G.D. Interpreting Neural-Network Connection Strengths. *AI Expert* **1991**, *6*, 46–51.
115. Lee, S.; Lee, K.-K.; Yoon, H. Using artificial neural network models for groundwater level forecasting and assessment of the relative impacts of influencing factors. *Hydrogeol. J.* **2019**, *27*, 567–579. [CrossRef]
116. Sharifi, Y.; Hosseinpour, M.; Moghbeli, A.; Sharifi, H. Lateral Torsional Buckling Capacity Assessment of Castellated Steel Beams Using Artificial Neural Networks. *Int. J. Steel Struct.* **2019**, *19*, 1408–1420. [CrossRef]
117. Clarke, S.; Robertson, I.T. A meta-analytic review of the Big Five personality factors and accident involvement in occupational and non-occupational settings. *J. Occup. Organ. Psychol.* **2005**, *78*, 355–376. [CrossRef]
118. Barrick, M.R.; Mount, M.K.; Li, N. The Theory of Purposeful Work Behavior: The Role of Personality, Higher-Order Goals, and Job Characteristics. *Acad. Manag. Rev.* **2013**, *38*, 132–153. [CrossRef]