ASCE

Prioritizing Safety Climate Improvements in the Indonesian Construction Industry Using Supervised Classification

Moslem Attar Raouf[1]; Riza Yosia Sunindijo, Ph.D.[2]; and Mojtaba Maghrebi, Ph.D.[3]

Abstract: Despite its significance, the Indonesian construction industry has poor safety performance. Improving the safety climate has been seen as a way to improve safety in the industry. Research on safety climate in this context has identified a range of safety climate factors that require improvements. However, construction organizations face difficulties in implementing improvement recommendations due to resource constraints. In order to help construction organizations in their efforts to improve the safety climate, this research demonstrates the use of supervised classification approaches to identify specific safety climate factors that construction organizations should focus on. Data were collected from 311 construction practitioners in Indonesia using a 22-item safety climate survey. Supervised classification methods, comprising ensemble methods, Support Vector Machine, Naïve Bayes, and Nearest Neighbor, were used. The analysis identified 14 safety climate items that can represent the original dataset with high accuracy (93%). These 14 items can be considered crucial items that should be prioritized in the Indonesian construction industry. These items revealed that, due to the high power distance culture in Indonesia, top-down approaches, such as giving clear instructions, providing training, and reminding people often about safety, are effective for engaging employees to focus on and participate in safety. The findings also suggest that understanding cultural context is important to determine effective strategies to improve safety. This research has also demonstrated the potential application of supervised classification approaches to help decision makers improve safety by focusing on crucial factors within a context. **DOI: 10.1061/JAEIED.AEENG-1588**. © *2023 American Society of Civil Engineers*.

Author keywords: Construction industry; Cultural relativity; Indonesia; Safety climate; Supervised classification.

Introduction

As the fourth most populous country in the world and one of the large emerging economies, Indonesia's economy has grown steadily in the past years. The Indonesian construction industry is an important sector that contributes to this economic growth. Supported by the commitment of the Indonesian government to improve the country's infrastructure, the growth of the Indonesian construction industry outpaces the national economic growth. The industry is also one of the largest employers, providing needed employment for a large number of the population, while contributing to the socioeconomic development of the nation (Lestari et al. 2020).

Despite its significance, the Indonesian construction industry has poor health and safety records, a fact that has often been acknowledged previously. As a result, research on health and safety in the context of the Indonesian construction industry has been increasing. One such research area focuses on safety climate, which refers to people's perceptions and attitudes toward health and safety in the workplace. Previously, safety climate research in the Indonesian construction industry has mainly focused on the dimensions and factor structure of safety climate, the relationship between safety climate and performance, the level of safety climate, the factors that influence safety climate, and recommendations to improve safety climate (Loosemore et al. 2019). This past research is important for revealing the safety climate status in the Indonesian construction industry, which has had little previous investigation.

Quantitative data collected using a questionnaire survey was the main method used in conducting safety climate research (Kalteh et al. 2021). These questionnaires consist of numerous items, which can be problematic for the Indonesian construction industry when prioritizing their attempt to improve safety climate. For instance, a framework identified 28 strategies for developing safety climate in the Indonesian construction industry (Lestari et al. 2020), which can be quite overwhelming for construction organizations and the industry when transforming themselves. There is a need, therefore, to operationalize safety climate by prioritizing important items that represent the shared perceptions of employees within a context (Zohar 2010). By building on these past efforts, this research aims to use an innovative data-driven approach to assist the Indonesian construction industry in improving its safety climate. Based on safety climate data, crucial factors that strongly influence safety climate are pinpointed, allowing construction organizations to focus on these factors in their efforts to improve safety climate.

Literature Review

Safety climate is a term coined by Zohar (1980, p. 96), who defined it as "a summary of molar perceptions that employees share about their work environments." The term *safety culture* was then proposed in a

[1]Dept. of Civil Engineering, Ferdowsi Univ. of Mashhad, Azadi Square, Mashhad, Razavi Khorasan Province 9177948974, Iran. ORCID: https://orcid.org/0000-0002-9203-9281. Email: mo.attarraouf@mail.um.ac.ir

[2]Associate Professor, School of Built Environment, UNSW Sydney (The University of New South Wales), Sydney NSW 2052, Australia (corresponding author). ORCID: https://orcid.org/0000-0001-5679-8937. Email: r.sunindijo@unsw.edu.au

[3]Associate Professor, Dept. of Civil Engineering, Ferdowsi Univ. of Mashhad, Azadi Square, Mashhad, Razavi Khorasan Province 9177948974, Iran. ORCID: https://orcid.org/0000-0001-7284-9095. Email: mojtabamaghrebi@um.ac.ir

Note. This manuscript was submitted on November 3, 2022; approved on March 27, 2023; published online on May 19, 2023. Discussion period open until October 19, 2023; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Architectural Engineering***, © ASCE, ISSN 1076-0431.

report by the INSAG (1986) after the Chernobyl nuclear power disaster, in which the report suggested that a lack of safety culture was one of the main causes of the disaster. In their following report, ACSNI (1993) defined safety culture as the product of individual and group values, attitudes, perceptions, competencies, and patterns of behavior that determine the commitment to, and the style and proficiency of, an organization's health and safety management.

At first, the terms *safety climate* and *safety culture* were used interchangeably, a situation that persists to a certain extent today (Kalteh et al. 2021). However, over the years, there is better clarity on what constitutes safety climate and safety culture. Safety climate is nowadays recognized as the psychological dimension of safety culture, reflecting the employees' perceptions and attitudes toward safety at a particular point in time. Safety climate is more dynamic, while safety culture tends to be more stable as it is linked to the embedded organizational culture (Loosemore et al. 2019).

Owing to the difficulties in assessing safety culture, safety climate has often been used as an indicator of safety performance, including that in the construction industry, to complement the typical lagging indicators, such as fatality and incidence rates (Kadir et al. 2022). Safety climate is sometimes considered a leading indicator due to its ability to assess performance before the occurrence of negative events (Zohar 2010). Furthermore, as an indicator of safety performance, safety climate has other advantages. It helps managers make investment decisions on areas that require improvements. When a safety climate survey is done regularly, it is useful for identifying trends in performance and establishing benchmarks. Collecting safety climate data is also relatively easy and economical, although it is important to note that safety climate cannot replace other measurement tools. Lastly, safety climate involves employees in the process and, when anonymity is ensured, employees are motivated to express their opinions without any fear of reprisal (Sunindijo and Zou 2012).

Owing to its advantages, the use of safety climate in construction safety research in Indonesia is growing. The existing literature reveals several distinct research areas on safety climate in the Indonesian construction industry. First, research has used safety climate to assess safety performance. Sunindijo et al. (2019) found that safety climate in building projects was higher than safety climate in a first-of-its-kind infrastructure project in Indonesia, even though the latter was managed by an international contractor with a high level of health and safety commitment. Loosemore et al. (2019) compared the level of safety climate in the Indonesian construction industry with that in the Australian construction industry and, surprisingly, found no significant difference.

Second, research focuses on gaining in-depth understanding of safety climate, particularly its dimensions and factor structure. Andi (2008) assessed six dimensions of safety climate, comprising top management commitment, safety rules and procedures, communication, worker competence, work environment, and worker involvement, in three types of construction projects and highlighted safety climate differences among the three projects. Based on qualitative feedback from respondents, Lestari et al. (2020) identified improvement gaps in the six dimensions of safety climate, comprising management commitment, communication, rules and procedures, supportive environment, personal accountability, and training.

Third, research investigates the impact of safety climate on project and organizational performance. Widjaja et al. (2020) found that safety climate is a factor that mediates the relationship between leadership and safety performance. Widyanty et al. (2020) found that safety climate has positive effects on construction organizations' competitive advantage and productivity and mediates the relationship between human resource management practice and competitive advantage.

Fourth, research finds factors that influence the level of safety climate. Kadir et al. (2022) investigated the influence of demographic characteristics on safety climate and suggested that attention should be given to employees who have a lower level of education, less work experience, and nonpermanent job status. Sunindijo et al. (2019) argued that higher project complexity hurts safety climate levels, while Loosemore et al. (2019) suggested that understanding cultural relativity is important when comparing safety climate across countries.

Table 1 summarizes previous research on safety climate in the Indonesian construction industry.

As presented in Table 1, safety climate surveys often contain many dimensions and items to capture the various manifestations of safety climate in practice, causing problems for construction stakeholders in prioritizing their improvement efforts. Focusing on particular items arbitrarily may result in redundant efforts if those items are not crucial in reality. There is a need to determine these crucial safety climate items in a more structured way to optimize efforts given limited resources. Contributing to enriching research in this area and facilitating the application of research findings to improve safety climate in the Indonesian construction industry, this research uses an innovative data-driven approach to eliminate safety climate items that do not contribute strongly to the actual level of safety climate.

Supervised Classification Approach

A two-step approach (Fig. 1) for supervised classification as introduced by Han et al. (2012) was developed and used in this research. The two steps are learning and classification (or class prediction).

In this approach, a model (or classifier) is trained in the learning step. Step 1 comprises three main components. The first component is preprocessing. Preliminary data cleaning, removal of unnecessary data attributes, and generating new attributes (such as class labels) are major tasks of this fundamental component. In the second component, the data instances are split into two sets of training and testing. Finally, in the third component, supervised learning algorithms are exploited and the model is trained. Since data tuples are labeled in multicategories, supervised learning algorithms are applied in this step.

With the model trained, the class of test data is predicted in the classification step (Step 2).

Case Study and Data Gathering

A 22-item safety climate questionnaire was used to assess five dimensions of safety climate, comprising management commitment, accountability, training, personal involvement, and engagement, in the Indonesian construction industry. A six-point Likert scale format, ranging from "strongly disagree" to "strongly agree" was used in the questionnaire. Using convenience sampling, the questionnaire was distributed to construction employees, including professionals, tradespeople, and laborers, in construction projects in Jakarta, the capital of Indonesia. In total, 311 valid responses were obtained. Table 2 presents the 22 items used in the questionnaire, along with their mean scores. The mode (the most frequently occurring value) for all of the items is 5. The bold items signify items that are retained at the end of the analysis in this paper. Factor analysis was performed and generated five dimensions of safety climate, which were subsequently named "management commitment," "awareness and accountability," "training," "personal involvement," and "engagement."

Table 1. Summary of safety climate in the Indonesian construction industry

Authors	Dimensions	No. of items	Findings
Andi (2008)	1. Top management commitment 2. Safety rules and procedures 3. Communication 4. Worker competence 5. Work environment 6. Worker involvement	31	Level of safety climate in three different types of construction projects was quite good. The levels of safety climate dimensions among the three projects were significantly different.
Loosemore et al. (2019)	1. Management commitment 2. Communication 3. Rules and procedures 4. Supportive environment 5. Personal accountability 6. Training	58	There is no difference in safety climate level between Australian and Indonesian construction industries. Cultural relativity should be considered when comparing safety climate across countries.
Sunindijo et al. (2019)	1. Management commitment 2. Communication 3. Rules and procedures 4. Supportive environment 5. Personal accountability 6. Training	58	Safety climate in building projects was higher than that in an infrastructure project. Project complexity has an influence on safety climate.
Lestari et al. (2020)	1. Management commitment 2. Communication 3. Rules and procedures 4. Supportive environment 5. Personal accountability 6. Training	58	A new integrated safety climate framework to improve safety performance in the Indonesian construction industry.
Widjaja et al. (2020)	Not reported	Not reported	Safety leadership has a positive influence on safety climate, which in turn has a positive influence on safety performance.
Widyanty et al. (2020)	1. Policy 2. Leadership 3. Participation 4. Communication	18	Human resource management practices have a positive influence on safety climate, which in turn has a positive influence on competitive advantage.
Kadir et al. (2022)	1. Management commitment 2. Priority of safety 3. Communication 4. OHS rules 5. Supportive environment 6. Involvement 7. Work environment 8. Personal priorities and need for safety 9. Personal appreciation of risk	55	There is a gap between safety management on paper and its actual implementation. Demographic characteristics influence the level of safety climate.

Methodology

A five-stage feature selection (FS) process (Fig. 2) introduced by Ang et al. (2016) was used. The main objective of the FS is to determine a subset of attributes such that they could describe the original dataset in the most accurate way possible. The outcome generated by this process is highly dependent on the decision made at each stage. An overview of the applied FS process and the performed stages are explained next.

Stage 1: Search Direction

In the first stage, the search direction and search starting point were determined. The starting point has a significant impact on the

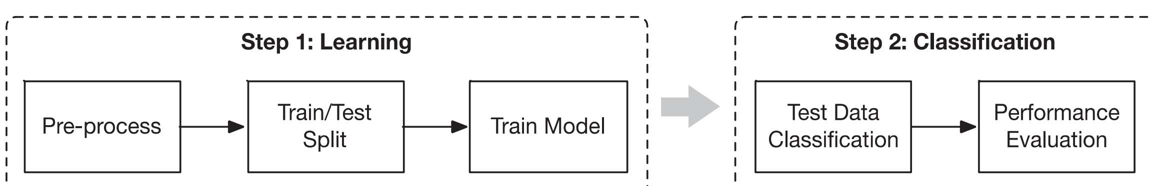

Fig. 1. Classification approach.

Table 2. Safety climate items and their corresponding means

No.	Item	Mean
A. Management commitment		
A1	**My project manager acts quickly to correct safety problems**	**5.022**
A2	**My direct supervisor pays attention to my safety**	**4.802**
A3	**My project manager expresses concern if safety procedures are not adhered to**	**4.363**
A4	**Safety procedures are carefully followed by all**	**5.009**
A5	**Safety communication is effective**	**4.980**
A6	Safety communication makes me pay attention on safety	5.113
A7	**Safety information is always brought to my attention by my direct supervisor**	**4.993**
A8	**My project manager is available for discussion when it comes to safety**	**4.821**
B. Awareness and accountability		
B1	I receive a lot of information about safety	4.958
B2	**Employees are always encouraged to focus on safety at their workplace**	**4.994**
B3	**A continuing emphasis on safety is important for me**	**5.003**
B4	**I understand all the safety rules**	**4.893**
B5	**Safety is the number one priority for me when completing a job**	**5.032**
B6	I am clear about my health and safety responsibilities	4.782
C. Training		
C1	**The company invests a lot of time and money in safety training**	**4.376**
C2	I am capable of identifying potentially hazardous situations	4.318
C3	The safety training provided is practical	4.886
D	Personal involvement	
D1	I can influence safety performance in my workplace	3.728
D2	No one criticizes me if I remind someone to work safely	4.699
E. Engagement		
E1	**I receive praise for working safely**	**4.374**
E2	**I am strongly encouraged to report unsafe conditions in my workplace**	**4.637**
E3	I am involved in implementing safety at work	4.624

Note: Bold items signify items that were retained at the end of the analysis in this paper.

search consequences and controls the search direction (Blum and Langley 1997). The two most common search directions are forward and backward. In forward search, the process starts with an empty subset, and at each iteration, a new feature is added to the subset. Backward search, in contrast, starts with the subset of all features and at each iteration, a redundant feature is removed from the subset. Another alternative is to start the search by randomly selected subset in the middle (Ang et al. 2016). Forward, backward, and random search directions are used in this paper.

Stage 2: Search Strategy

There are three categories of search strategies: exponential, sequential, and randomized. Exponential (or complete) search is the most computational exhaustive because all possibilities are calculated and then the optimal subset is determined. This strategy is impractical for FS over large datasets (Ang et al. 2016).

In a sequential search, a feature is added to or removed from the subset considering its value at each iteration. The impact of adding or removing any feature to the overall accuracy of the model is interpreted as the value of the feature. The three most-common sequential search strategies mentioned in the literature are: forward sequential selection (FSS), backward sequential selection (BSS), and hill climbing (HC). FSS (Fig. 3) aims to find the best feature with highest value at each iteration, which is then added to an originally empty set of optimal features. In BSS (Fig. 4), the original subset of optimal features includes all features. In this approach, the most redundant feature with lowest value is identified and removed at each iteration.

Both FSS and BSS are among greedy sequential search strategies (Khaire and Dhanalakshmi 2022). FSS and BSS do not yield similar results due to the direction of search sequence. In addition, depending on the designated target number of features, one could be faster than the other (Pedregosa et al. 2011). In HC, a feature is added or removed from the dataset at an iteration. The search is performed for finding the optimal features from a random set of features, which is then followed by inverting the current status of each feature in the subset. The random nature of HC increases its complexity (Khaire and Dhanalakshmi 2022), and therefore was not selected in this paper.

In a randomized search, a random subset of features is selected and then the search proceeds in either sequential or random order, which means the search will perform with no regular movement. Heuristic search algorithms, such as genetic algorithm (GA) or tabu search, are typically used in this strategy (Ang et al. 2016).

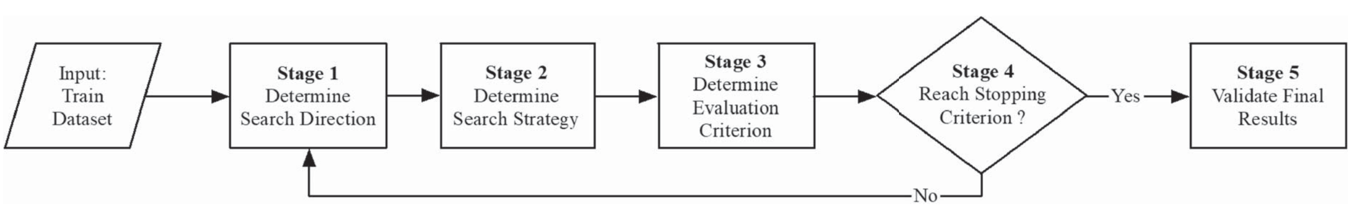

Fig. 2. FS process.

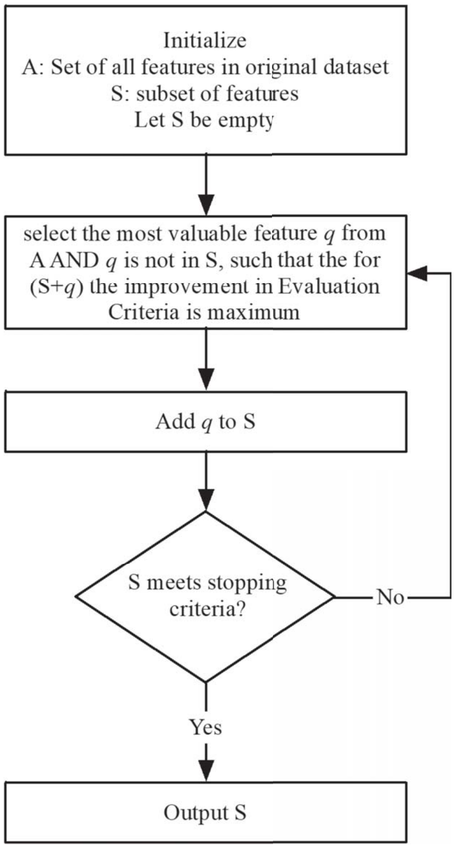

Fig. 3. Forward sequential search.

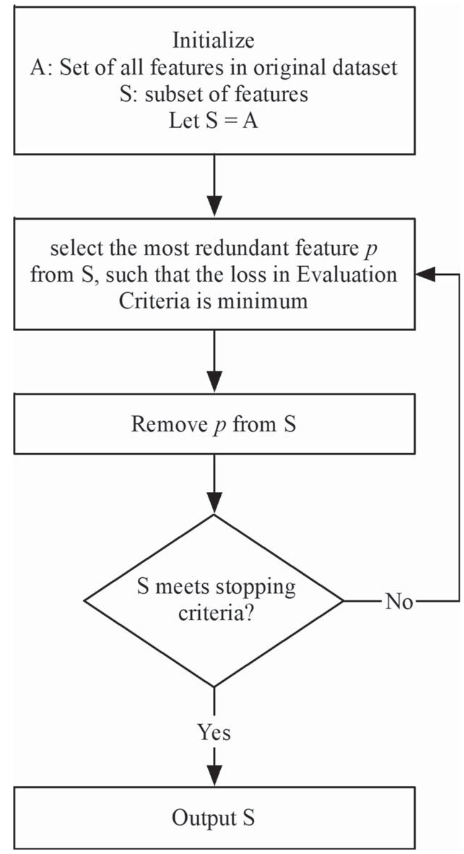

Fig. 4. Backward sequential search.

GA (Fig. 5), which is among the population methods of optimization, is inspired by biological evolution, where fitter individuals are more likely to pass on their genes to the next generation (Kochenderfer and Wheeler 2019). At first, initial solutions are randomly generated in GA. Then at each iteration, there is a chance between two solutions to swap their components. Next, the solutions are evaluated against an evaluation fitness function (GA-objective). Solutions with better evaluation results would have a higher chance of being regenerated. Unless any of the stopping criteria is met, a new generation is created and crossover and mutation take place as well. The population is replaced with the newly generated population and the next iteration is initiated (Kochenderfer and Wheeler 2019). Examples of stopping criteria begin to reach the maximum limits of iterations, reaching to the maximum number of iterations without any improvement, or obtaining the desired result with minimum requirements.

Stage 3: Evaluation Criteria

The main objective of the FS is to determine a subset of attributes such that they could describe the original dataset in the most possible accurate way. There are four categories of FS techniques introduced in the literature: filter, wrapper, embedded, and hybrid methods.

First, filter methods rely on characteristics of data to assess feature importance and are independent of any learning algorithms. Filter methods are typically more computationally efficient than wrapper methods, but it may not provide the best results because there is no specific learning algorithm to guide the FS phase (Li et al. 2018).

Second, wrapper methods evaluate the quality of feature subsets using the results of a specific learning algorithm typically in two steps: (1) search and determine a subset of features, and (2) evaluate

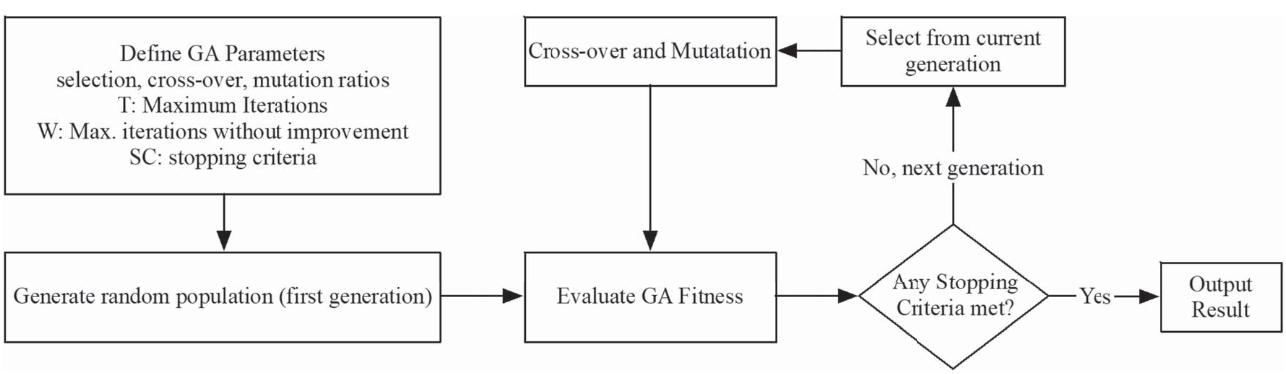

Fig. 5. Genetic algorithm.

the subset. These two steps are repeated until some predetermined stopping criteria are met. Examples of stopping criteria could be the "minimum acceptable performance rate" or "minimum acceptable number of selected features." The main disadvantage of wrapper methods is that they typically have a higher computational cost. The search space for learning algorithm is 2^d for d features, which makes wrapper methods impractical when d is very large. To address this, heuristic search methods such as GA could be applied (Khaire and Dhanalakshmi 2022; Li et al. 2018).

Third, in embedded methods, a classifier is developed, and then a search is performed for determining ideal feature subsets. While the advantages of embedded methods are similar to wrapper methods, their computational cost is better than wrapper methods (Khaire and Dhanalakshmi 2022).

Recently some literature added hybrid methods to the previous three common methods. Hybrid methods attempt to benefit from the advantages of filter and wrapper methods. They aim to have a good compromise between efficiency (computational effort) and effectiveness (quality in the associated objective task when using the selected features) (Li et al. 2018; Solorio-Fernández et al. 2020).

From various categories of supervised learning procedures, six algorithms were selected and applied (Table 3). All classifiers used in this research were developed by Scikit-Learn (Pedregosa et al. 2011).

Ensemble Methods

The core of ensemble learning models is to combine predictions from a number of learners. The ensemble methods should be far more accurate than the individual classifiers since they leverage the aggregation of different classifier outputs (Nzuva and Nderu 2019). Sagi and Rokach (2018) identify several challenges in machine learning that could be mitigated using ensemble methods. These challenges are class imbalance, concept drift, and curse of dimensionality. Class imbalance is when one class has substantially more frequency than other classes. Concept drift refers to when the distribution of features and the labels tend to change over time. Curse of dimensionality is caused by an increase in features fed to the model that expands the search space exponentially and lowers the fitting probability of the model.

Ensemble methods are generally categorized into two groups: averaging and boosting. The key principle in averaging is to build several estimators individually and then to average the predicted outcome. Examples of this group are bagging and random forests methods. In boosting, however, base estimators are built consequently with the aim that each one reduces the bias of the aggregated estimator. AdaBoost and gradient tree boosting are among examples of boosting group (Pedregosa et al. 2011).

Following ensemble methods were selected and applied for FS in this paper:

1. Random forests (RF), or random decision forests, are an ensemble of learning methods for classification. During the training model, multiple decision trees are constructed (hence the name forest). The classification task is completed when the result is generated by a majority of the decision trees for any given parameter(s) (Breiman 2001). RF is a popular ensemble method mainly due to its simplicity and predictive performance. Moreover, it is easier to adjust the RF model than other methods (Sagi and Rokach 2018).
2. Bagging, also known as bootstrap aggregation, is an ensemble method developed to enhance the accuracy and stability of supervised classification algorithms. The bagging classifier model requires developing numerous instances of black-box estimators (or inducers) based on random subsets of the original training set (i.e., samples). The model then aggregates their individual predictions to determine how a new instance is classified (Nzuva and Nderu 2019; Pedregosa et al. 2011).

   Sagi and Rokach (2018, sec. 4.2) declare that "since sampling is done with replacements, some of the original instances are likely to appear more than once when training the same inducer while other instances may not be included at all. Since the inducers are independently trained, bagging can easily be implemented in a parallel manner by training each inducer using different computational units."
3. AdaBoost was introduced by Freund and Schapire (1997). The main characteristic of AdaBoost is that it has no random elements but its core principle is to fit a sequence of weak learners, such as small decision trees, on successive reweightings of the training set. Weak learners are models that are only slightly better than random guessing. The final prediction is delivered by a combination of all of weak learners through a weighted majority vote (Breiman 2001; Pedregosa et al. 2011).

Table 3. Applied supervised classification algorithms

Category of supervised classification	Applied algorithm
Ensemble methods	1. RF 2. Bagging 3. Adaboost
SVM method	4. Linear SVM
Naïve Bayes methods	5. GNB
Nearest Neighbor methods	6. KNN

Support Vector Machine

Support vector machine (SVM) is a classification method for both linear and nonlinear data. "In a nutshell, an SVM is an algorithm that uses a nonlinear mapping to transform the original training data into a higher dimension. Within this new dimension, it searches for the linear optimal separating hyperplane, i.e., a 'decision boundary' separating the tuples of one class from another. With an appropriate nonlinear mapping to a sufficiently high dimension, data from two classes can always be separated by a hyperplane. The SVM finds this hyperplane using support vectors ('essential' training tuples) and margins (defined by the support vectors)" (Han et al. 2012, p. 408). In addition to classification, SVMs are also used for regression and outlier detection. SVM methods are effective in high dimensional spaces and are also memory efficient by using a subset of training points in the decision function. The SVM classifier used in this research is versatile, with different kernel functions that can be specified for the decision function (Pedregosa et al. 2011).

Naïve Bayes

"Naive Bayes [NB] is the simplest form of Bayesian network, in which all attributes are independent given the value of the class variable" (Zhang 2004, para. 5). The fundamental element of this set of supervised learning methods is applying Bayes' theorem with the "naïve" assumption. Bayes' theorem is a way to calculate posterior probability, which can be represented as

$$P(H|\boldsymbol{X}) = \frac{P(\boldsymbol{X}|H)P(H)}{P(\boldsymbol{X})} \tag{1}$$

where $P(H|\boldsymbol{X})$ = posterior probability, of H conditioned on $\boldsymbol{X}$. NB (or simple Bayesian) classifier works as follows (Han et al. 2012; Pedregosa et al. 2011; Zhang 2004). Let $\boldsymbol{X}$ represent a training tuple vector with n-dimension attributes, $\boldsymbol{X} = (x_1, x_2, \ldots, x_n)$. Suppose that there are m classes, $C_1, C_2, \ldots, C_m$. By using the Bayes' theorem (1) the posterior probability of tuple $\boldsymbol{X}$ to be

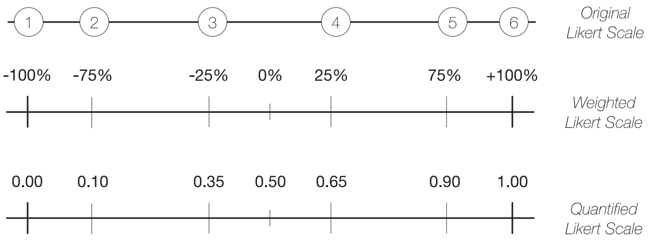

Fig. 6. Quantify Likert scale diagram.

class y, we have

$$P(y|\boldsymbol{X}) = \frac{P(\boldsymbol{X}|y)P(y)}{P(\boldsymbol{X})} \tag{2}$$

where $y =$ class variable; and $P(y|\boldsymbol{X}) =$ NB classifier and predicts that tuple $\boldsymbol{X}$ belongs to the Class y if and only if

$$P(y|\boldsymbol{X}) > P(c_j|\boldsymbol{X}) \quad \text{for } 1 \leq j \leq m, y \neq c_j \tag{3}$$

This means that the predicted class is y for which $P(\boldsymbol{X}|y)P(y)$ is the maximum. With the *naïve* assumption that all attributes are independent given y, that is

$$P(\boldsymbol{X}|y) = P(x_1, x_2, \ldots, x_n|y) = \prod_{k=1}^{n} P(x_k|y) \tag{4}$$

The resulting classifier is then

$$P(y|\boldsymbol{X}) = \frac{\prod_{k=1}^{n} P(x_k|y)P(y)}{P(\boldsymbol{X})} \tag{5}$$

Since $P(\boldsymbol{X})$ is constant, the NB classifier will be

$$P(y|\boldsymbol{X}) \propto P(y)\prod_{k=1}^{n} P(x_k|y) \Rightarrow \hat{y} = \arg\max_{y} P(y)\prod_{k=1}^{n} P(x_k|y) \tag{6}$$

with a different assumption about the distribution of $P(x_k|y)$, different NB classifiers can be provided. For instance, in the GNB classifier, the likelihood of the features is assumed to be Gaussian, which can be shown as

$$P(x_k|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_k - \mu_y)^2}{2\sigma_y^2}\right) \tag{7}$$

The parameters σ_y and μ_y are estimated using maximum likelihood (Pedregosa et al. 2011).

Although NB classifiers are based on over-simplified assumptions, they have relatively exceptional performance in real-world classification problems such. As compared with other more complex methods, NB models are extremely fast, and they can cope with the curse of dimensionality (Pedregosa et al. 2011). Zhang (2004) explains the characteristics of NB classifiers and discusses theoretical reasons on why they work well and with which types of data it does.

Nearest Neighbors

Nearest neighbors are among lazy learning techniques in classification. Nearest neighbor models are capable of performing both supervised and unsupervised learning. Upon receiving the training data, the lazy model simply stores it with minor processing and waits until receiving the test tuple. The model classifies the test data only when it receives it. The classification is based on the similarity of testing tuple to the stored training data. In contrast to lazy learner models, eager learners construct a classification model upon receiving the training tuples. In other words, in an eager learner model, the classifier is developed, ready, and eager to classify new (i.e., test) tuples (Han et al. 2012). Since neighbor-based methods are for simply remembering all of their training data, they are also known and referred to as nongeneralizing machine learning methods (Pedregosa et al. 2011).

For this research, the K-nearest neighbor (KNN) ($K = 5$) method was selected and applied. "When given an unknown tuple, a k-nearest-neighbor classifier searches the pattern space for the k training tuples that are closest to the unknown tuple. These k training tuples are the k 'nearest neighbors' of the unknown tuple' (Han et al. 2012, p. 423).

Stage 4: Stopping Criteria

The search process stops once any of the stopping criteria is met. Some of the common stopping criteria are:

1. Target number of features.
2. Number of iterations.
3. Rate of improvements between two consecutive iterations.
4. Obtaining a subset of feature according to a defined evaluation function.

In this research, the FS process is performed in three rounds. The target number of features was set to 14, 11, and 6, respectively, in each round. Moreover, the GA for random search halted if there is no improvement in 30 consecutive iterations.

Stage 5: Result Validation

Results obtained by the search process are validated in the last stage. The most common validation techniques are cross validation (CV) and performance evaluation using confusion matrix. For this research, "accuracy measurement" was selected as the evaluation criterion for FS.

Accuracy, being one of the most popular metrics in multiclass classification, returns an overall measure of how much the model

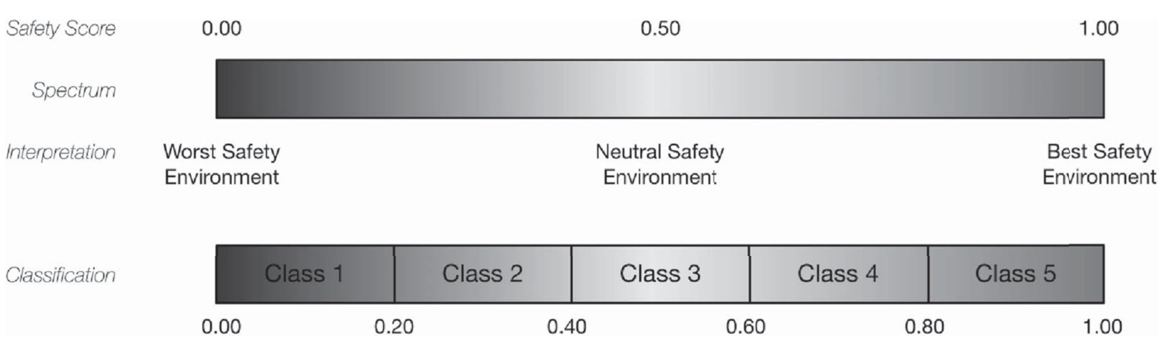

Fig. 7. Spectrum of safety score and uniformed-sized classes.

correctly predicts the entire set of data, and it is directly computed from the confusion matrix using (Grandini et al. 2020)

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{8}$$

True Positive (TP) and True Negative (TN) in the numerator are the entries classified correctly by the classifier. The denominator is the sum of all entries of the confusion matrix, which includes all correct and incorrect classifications.

Results

The original dataset consists of 22 six-point Likert scale responses collected from 311 respondents. Some researchers consider Likert scale as an ordinal attribute (Norman 2010; Schumacker and Lomax 2015). Han et al. (2012, p. 42) defines ordinal attribute as "an attribute with possible values that have a meaningful order or ranking among them, but the magnitude between successive values is not known." Moreover, the use of various parametric methods, such as analysis of variance,

Table 4. Results of all applied classifiers

Search parameters	Classifier ID	Stopping criteria	Subset	Accuracy	Training time (sec.)
BSS	M1-RF	14	00001011-111011-110-01-111	89.36%	21.990
FSS			11011011-101011-101-00-110	92.55%	29.928
GA_Search			11010010-110111-101-10-110	92.55%	32.044
BSS	M2-Bagging		10010000-011110-111-11-111	88.30%	4.040
FSS			11111011-110011-100-00-110	91.49%	4.720
GA_Search			10110111-001011-100-10-111	91.49%	9.822
BSS	M3-AdaBoost		00000000-111111-111-11-111	84.04%	11.654
FSS			11111111-111110-000-00-001	84.04%	16.042
GA_Search			01011011-100011-101-10-111	84.04%	15.743
BSS	M4-SVM		01001000-110101-111-11-111	85.11%	1.042
FSS			11011111-111001-100-10-001	89.36%	1.048
GA_Search			01010111-101010-010-11-111	91.49%	4.213
BSS	M5-GNB		00010001-111010-111-11-111	84.04%	0.594
FSS			01101111-110110-100-00-111	91.49%	0.798
GA_Search			01000110-011111-011-10-111	90.43%	2.006
BSS	M6-KNN		10000011-101011-111-01-111	90.43%	1.119
FSS			11111011-011110-100-00-110	93.62%	1.333
GA_Search			10001101-111011-101-10-110	93.62%	3.830
BSS	M1-RF	11	00000001-101011-110-01-111	89.36%	27.538
FSS			11001001-001011-101-00-110	90.43%	26.018
GA_Search			10100101-000100-111-00-111	89.36%	28.493
BSS	M2-Bagging		00010000-001110-101-11-111	88.30%	4.609
FSS			11111011-010000-100-00-110	92.55%	4.075
GA_Search			01101010-101110-110-00-100	90.43%	6.416
BSS	M3-AdaBoost		00000000-000111-111-11-111	84.04%	14.743
FSS			11111111-011000-000-00-001	84.04%	13.529
GA_Search			00101111-011110-010-00-001	84.04%	14.803
BSS	M4-SVM		01001000-000101-011-11-111	85.11%	1.262
FSS			11011111-101001-000-00-001	89.36%	0.874
GA_Search			10101100-010011-000-10-111	89.36%	3.395
BSS	M5-GNB		00000001-111000-111-11-101	84.04%	0.746
FSS			01100011-110010-100-00-111	92.55%	0.739
GA_Search			00000011-111100-111-00-110	92.55%	1.998
BSS	M6-KNN		10000000-101011-111-01-110	90.43%	1.431
FSS			11111011-010100-000-00-110	91.49%	1.137
GA_Search			11000110-111010-010-00-011	90.43%	3.796
BSS	M1-RF	6	00000000-000000-110-01-111	89.36%	34.578
FSS			11000001-000011-000-00-100	89.36%	15.672
GA_Search			00011010-101000-000-10-000	88.30%	40.454
BSS	M2-Bagging		00010000-000110-000-00-111	89.36%	5.722
FSS			11100001-000000-000-00-110	90.43%	2.483
GA_Search			00111001-000000-000-00-110	91.49%	7.650
BSS	M3-AdaBoost		00000000-000000-001-11-111	84.04%	18.355
FSS			11111000-001000-000-00-000	81.91%	8.298
GA_Search			00010011-100010-000-00-001	85.11%	14.047
BSS	M4-SVM		01000000-000000-000-11-111	85.11%	1.558
FSS			01010111-000000-000-00-001	88.30%	0.516
GA_Search			11000010-010000-000-00-011	88.30%	3.574
SS	M5-GNB		00000000-100000-110-11-100	80.85%	0.947
FSS			01000010-100000-100-00-110	91.49%	0.422
GA_Search			00000010-110000-100-00-011	88.30%	1.895
BSS	M6-KNN		10000000-100000-100-01-110	87.23%	1.649
FSS			11100010-000100-000-00-010	91.49%	0.690
GA_Search			00010001-100001-000-00-110	90.43%	4.363

Table 5. Summary of results, accuracy

	Search strategy			
Evaluation criteria	BSS	FSS	GA_Search	Overall average accuracy
14 Features	86.88%	90.43%	90.60%	89.30%
M1-RF	89.36%	92.55%	92.55%	91.49%
M2-Bagging	88.30%	91.49%	91.49%	90.43%
M3-AdaBoost	84.04%	84.04%	84.04%	84.04%
M4-SVM	85.11%	89.36%	91.49%	88.65%
M5-GNB	84.04%	91.49%	90.43%	88.65%
M6-KNN	90.43%	93.62%	93.62%	92.56%
11 Features	86.88%	90.07%	89.36%	88.77%
M1-RF	89.36%	90.43%	89.36%	89.72%
M2-Bagging	88.30%	92.55%	90.43%	90.43%
M3-AdaBoost	84.04%	84.04%	84.04%	84.04%
M4-SVM	85.11%	89.36%	89.36%	87.94%
M5-GNB	84.04%	92.55%	92.55%	89.71%
M6-KNN	90.43%	91.49%	90.43%	90.78%
6 Features	85.99%	88.83%	88.66%	87.83%
M1-RF	89.36%	89.36%	88.30%	89.01%
M2-Bagging	89.36%	90.43%	91.49%	90.43%
M3-AdaBoost	84.04%	81.91%	85.11%	83.69%
M4-SVM	85.11%	88.30%	88.30%	87.24%
M5-GNB	80.85%	91.49%	88.30%	86.88%
M6-KNN	87.23%	91.49%	90.43%	89.72%
Overall average accuracy	86.58%	89.78%	89.54%	88.63%

regression, and correlation, are not recommended for Likert scale data (Norman 2010).

The use of arithmetic mean and standard deviation for research or analysis on the ordinal data are also not recommended (Han et al. 2012; Stevens 1946) for, as Stevens stated in 1946, "these statistics imply a knowledge of something more than the relative rank-order of data." For measuring the central tendency of ordinal data, the mode and median should be used (Han et al. 2012). This research adopts this view, although we acknowledge that there is an opposing view that advocates the use of parametric analyses of Likert scale data.

The Likert scale in the original questionnaire is assumed as bipolar. The minimum and maximum quantified values of the Likert scale in this research were considered as 0 and 1, respectively. In addition, a different weight was assigned to each Likert scale. The reason for this was to develop a model in which the difference between "strongly agree (6)" and "agree (5)" would be less than "slightly agree (4)" and "slightly disagree (3)." Fig. 6 presents the Likert scale corresponding values, weight, and quantified score.

To obtain the safety dimension, the median of safety features associated with each dimension was calculated and then quantified using the assigned scales presented in Fig. 6. The dimension vector for tuple x is $D_x = (d_{1x}, d_{2x}, d_{3x}, d_{4x}, d_{5x})$, where $d_{ix} = i$th dimension.

Table 6. Summary of results, elapsed training time

	Search strategy			
Evaluation criteria	BSS	FSS	GA_Search	Average training time
M1-RF	28.04	23.87	33.66	28.52
M2-Bagging	4.79	3.76	7.96	5.50
M3-AdaBoost	14.92	12.62	14.86	14.13
M4-SVM	1.29	0.81	3.73	1.94
M5-GNB	0.76	0.65	1.97	1.13
M6-KNN	1.40	1.05	4.00	2.15
Average training time	8.53	7.13	11.03	8.90

The safety score for each respondent, using simple additive weighting (SAW), would be equal to the sum product of the "weight" and "dimension" vectors, as presented in

$$\text{Safety Score}_x = \sum_{i=1}^{5} (w_i \times d_{ix}) \tag{9}$$

where the weight vector W = importance of each safety dimension by the number of features it has. The weight shall be a vector in the form $W = (w_1, w_2, \ldots, w_5)$ such that $\sum_{i=1}^{5} w_i = 1$. The weight reflects the importance of each relative dimension to the overall safety score. The weight vector was determined using

$$w_i = \frac{\text{Number of attributes in Dimension } i}{\text{Total number of attributes}}, \quad i = \{1,\ 2,\ 3,\ 4,\ 5\} \tag{10}$$

The safety score is a float number from 0.00 to 1.00, representing lowest (or worse) and highest (or best) safety climate, respectively. Finally, the safety score was classified in five safety classes as presented in Fig. 7.

The training dataset was used for training all models and consisted of 70% of the original dataset tuples (217 of 311). The remaining 94 tuples were considered as the test dataset to measure the performance of the applied model. Train/test data tuples were selected randomly, and no test data was provided for training models.

The three applied search strategies, described in the Methodology section, are: (1) forward sequential search; (2) backward sequential search; and (3) random start point with GA. The GA model developed by Ryan (Mohammad) Solgi was used in this research (Solgi, n.d.).

Embedded methods, with training supervised classifier, were applied for FS. Supervised classification is applicable because data will be labeled with the safety class. The selected classifier shall also be able to perform multiclass classification.

FS was performed in three rounds, each for finding best 14, 11, and 6 features (i.e., stopping criteria). Using the "search strategies" and "algorithms" detailed in the Methodology section, the models are trained and then generate a subset of features with an identified targeted number of features. "Test data" will then fed to the trained model to predict the safety classes. The accuracy measure is used to evaluate the performance of the model. The resulting selected features and their relevant accuracy score are presented in Table 4.

In Table 4, the selected search parameter is provided in the first column (Stages 1 & 2). BSS and FSS represent backward sequential search and forward sequential search, respectively. "Rand. GA" represents random starting point with random search direction, and GA is used for this search method. "Classifier ID" contains the id/name of the applied algorithm (Stage 3) and stopping criteria (Stage 4) and is self-explanatory. The subset is a 22-character-long string consisting of binary values of "0" (exclusion) and "1" (inclusion) of the relevant feature in the subset. Features in each dimension are separated with dash ("-") character. For each obtained subset, the relevant accuracy score and elapsed time for training the model are also provided in the results table (Table 4).

Accuracy and elapsed training time were evaluated to validate and select the best results. Accuracy, as the main objective of the process, represents the performance of the model and elapsed training time provides how fast the results yielded in each model.

As expected with reducing number of target features, the accuracy decreases on average as well. The average accuracy for selecting 14, 11, and 6 features are 89.30%, 88.77%, and 87.83%. Regarding the search parameters, FSS and random search with

J. Archit. Eng., 2023, 29(3): 04023016

Table 7. Final results

	Search parameters		
Item	FSS	FSS	FSS
Classifier ID	M6-KNN	M5-GNB	M5-GNB
Stopping criteria	14	11	6
A1	1	0	0
A2	1	1	1
A3	1	1	0
A4	1	0	0
A5	1	0	0
A6	0	0	0
A7	1	1	1
A8	1	1	0
B1	0	1	1
B2	1	1	0
B3	1	0	0
B4	1	0	0
B5	1	1	0
B6	0	0	0
C1	1	1	1
C2	0	0	0
C3	0	0	0
D1	0	0	0
D2	0	0	0
E1	1	1	1
E2	1	1	1
E3	0	1	0
Accuracy	93.62%	92.55%	91.49%
Training time (sec.)	1.333	0.739	0.422

GA yielded higher accuracy results with an overall average accuracy of 89.78% and 89.54%, respectively. BSS produced the lowest average accuracy. Considering the evaluation criteria (applied classifiers in Stage 3), KNN and bagging have the best average accuracy score overall. Tables 5 and 6 present these results.

Considering the elapsed time to train the model, FSS was the fastest search strategy. As for the evaluation criteria, GNB has the lowest elapsed training time on average.

For selecting the final results for each stopping criteria, higher accuracy and lower training time were considered as primary and secondary selection parameters. Considering this, the final three models are extracted and presented in Table 7. Features A2, A7, C1, E1, and E2 are selected repeatedly in each round of FS.

The confusion matrix for multiclass classification as described by Tharwat (2021) was used to validate selected results. All classifiers performed well for predicting tuples with safety class = 5, with a precision rate of above 94%. For testing tuples with actual safety class = 4, KNN has the best precision over GNB with a precision of 90%. In addition to precision, recall and f1-score of the classifiers results presented in Fig. 8.

Discussion

The research has used supervised classification approaches to identify important safety climate items that Indonesian construction organizations should focus on to improve their safety performance. Here, KNN and GNB seem to be the best classification algorithms, generating solutions with the highest accuracy. Considering the results presented in Table 7, selecting 14 safety climate items is the optimal solution since in addition to accuracy, it has the highest prediction score as well. Other two results, however, have acceptable and close accuracy. In practice, organizations can select the expected standard of accuracy to identify the minimum number of items to be used in decision making. These 14 safety climate items were bolded in Table 2.

The most glaring result is the removal of the personal involvement dimension. Considering the context, this is actually understandable. In Indonesia as a country with a high power distance culture (Loosemore et al. 2019), those with a high power distance orientation prefer directive leaders and expect to receive clear directions and instructions from their managers. In fact, breaking from hierarchical leadership styles and delegating more responsibly and autonomy to subordinates can have negative consequences on satisfaction for those high on power distance (Daniels and Greguras 2014). Encouraging the involvement of employees, therefore,

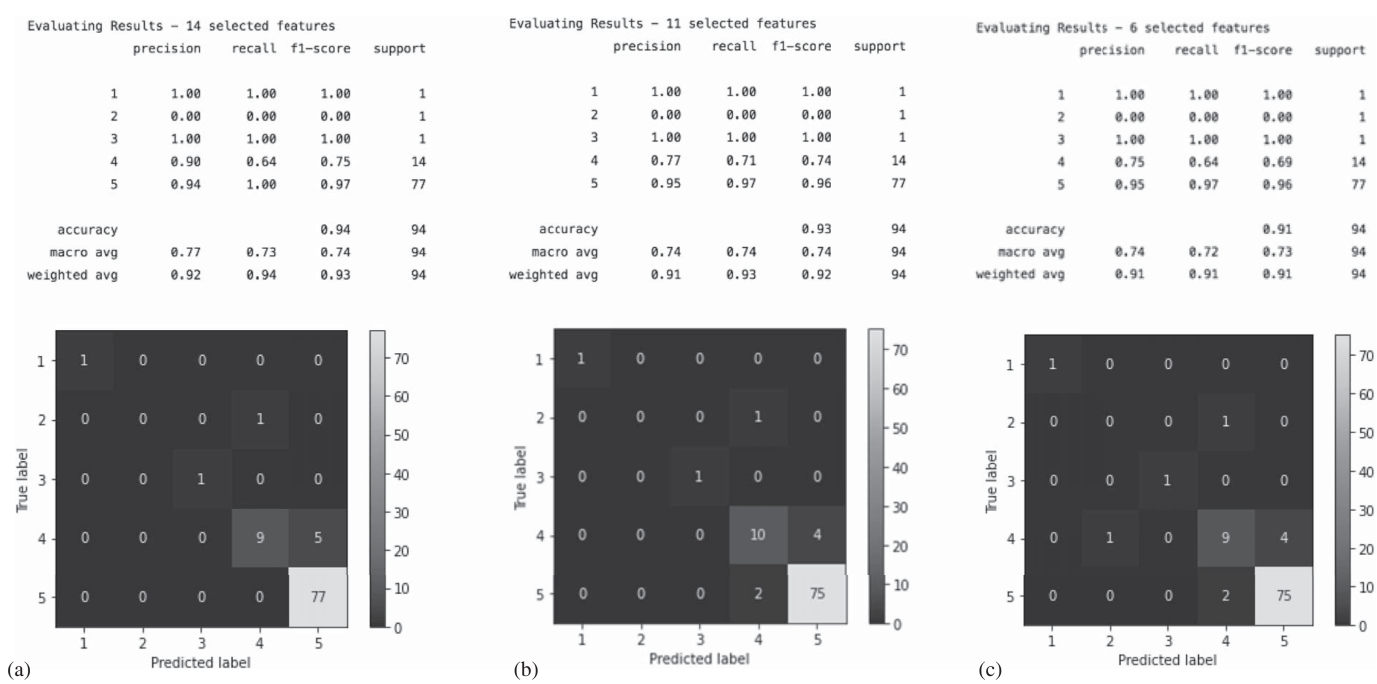

Fig. 8. Classification report and confusion matrix for (a) 14; (b) 11; and (c) 6.

although this certainly has merit, would not be as effective as providing clear top-down instructions on safety. Furthermore, many workers in the Indonesian construction industry have a low level of safety knowledge and awareness due to the informality of the sector (Chan and Pribadi 2022). This can limit their safety involvement and bottom-up feedback through engagement (Yip et al. 2012) as these processes require an adequate level of safety knowledge.

Although this dimension is excluded, it is important to note that involving and engaging employees is still important for safety. However, cultural relativity should be considered when engaging and involving employees in the Indonesian construction industry, and our method is useful for identifying the appropriate approach. For instance, within the awareness and accountability dimension, it is crucial to remind employees frequently to focus on safety (B2) and to ensure that employees understand all the safety rules (B4). The training dimension indicates that providing safety training is a way to engage employees and improve their safety knowledge (C1). Giving praise or reward for working safely (E1) seems to be another way to engage employees effectively, and there is a need to do this more given the relatively low mean score. It is clear that these approaches lean toward top-down approaches, which are appropriate considering the high power distance culture in Indonesia. Through these approaches, employees can then be involved in safety implementation by equipping them to report unsafe conditions in the workplace (E2) and making safety a priority in completing tasks (B5).

Finally, management commitment is a crucial dimension, which again indicates a top-down approach to promote a safety climate. Previous research in Indonesia found that managers in construction organizations are inconsistent with their attention to safety and that there is a gap between what they say and their action (Kadir et al. 2022). Often safety is being highlighted and communicated but its application is fairly limited usually due to the drive to reduce cost. This reality, into a certain extent, is shown by the lack of investment in safety training (C1), which has a relatively low safety climate score. There is also a need for project managers to express safety concerns when safety implementation is below standard (A3), which is particularly important given their as role as leaders at the project level.

Conclusions

Improving safety climate has been accepted as a way to improve safety performance in the Indonesian construction industry. By using a supervision classification approach, this research has simplified the process of developing safety climate by focusing on safety climate items that matter the most. This is important so that construction organizations can use their limited resources intelligently to implement strategies to develop specific safety climate items that can have the most influence on safety climate in the Indonesian construction industry.

Using the approach proposed in this research, 22 safety climate items can be reduced to 14 items while maintaining a high level of accuracy (93.62%) to represent the overall safety climate. As such, theoretically, this research has introduced a new approach to reduce questionnaire length while maintaining a high level of accuracy to ensure that the shortened questionnaire still represents the construct being measured. Practically, the results reveal that, in terms of involving and engaging employees in the Indonesian construction industry, top-down approaches are effective due to the country's high power distance and low safety knowledge among their construction workforce. This result shows the importance of considering cultural relativity when developing strategies to improve safety. Rather than blindly adopting safety measures implemented in other countries (usually from developed countries with Western culture), it is important to tailor improvement strategies that align with the cultures and norms in Indonesia to ensure success.

There are several research limitations worth noting. The first is related to the calculation of the safety score and the assignment of safety classes. SAW was used in this research to convert five safety dimensions to one scalar metric. It is expected that classifiers yield different results if other techniques such as Euclidean or Manhattan distance were applied. Another limitation is related to performing GA with a random start point, since the results provided by this search strategy are not constant. Moreover, as presented in Table 4 and expressed by Ang et al. (2016), results are classifier specific. There is almost no consistent pattern among the features selected by various classifiers. Finally, FS performed with accuracy measure as the single evaluation criterion. In future research, additional metrics, such as consistency between selected features can be assessed.

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request. The Jupyter Notebook Code is available online at https://github.com/moslem-raouf/safety_climate_features.

References

ACSNI (Advisory Committee on the Safety of Nuclear Installations). 1993. *ACSNI study group on human factors. 3rd report: Organising for safety*. London: HM Stationery Office.

Andi. 2008. "Construction workers perceptions toward safety culture." *Civ. Eng. Dimension* 10 (1): 1–6.

Ang, J. C., A. Mirzal, H. Haron, and H. N. A. Hamed. 2016. "Supervised, unsupervised, and semi-supervised feature selection: A review on gene selection." *IEEE/ACM Trans. Comput. Biol. Bioinf.* 13 (5): 971–989. https://doi.org/10.1109/TCBB.2015.2478454.

Blum, A. L., and P. Langley. 1997. "Selection of relevant features and examples in machine learning." *Artif. Intell.* 97 (1): 245–271. https://doi.org/10.1016/S0004-3702(97)00063-5.

Breiman, L. 2001. "Random forests." *Mach. Learn.* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Chan, T.-K., and K. S. Pribadi. 2022. *Construction in Indonesia looking back and moving forward*. 1st ed. London: Routledge.

Daniels, M. A., and G. J. Greguras. 2014. "Exploring the nature of power distance: Implications for micro- and macro-level theories, processes, and outcomes." *J. Manage.* 40 (5): 1202–1229. https://doi.org/10.1177/0149206314527131.

Freund, Y., and R. E. Schapire. 1997. "A decision-theoretic generalization of on-line learning and an application to boosting." *J. Comput. Syst. Sci.* 55 (1): 119–139. https://doi.org/10.1006/jcss.1997.1504.

Grandini, M., E. Bagli, and G. Visani. 2020. *Metrics for multi-class classification: An overview*.

Han, J., M. Kamber, and J. Pei. 2012. *Data mining concepts and techniques*. Burlington, MA: Morgan Kaufmann Publishers.

INSAG (International Nuclear Safety Advisory Group). 1986. *Summary report on the post-accident review meeting on the chernobyl accident*. Vienna, Austria: International Atomic Energy Agency.

Kadir, A., F. Lestari, R. Y. Sunindijo, D. Erwandi, Y. Kusminanti, R. Modjo, B. Widanarko, and N. A. Ramadhan. 2022. "Safety climate in the Indonesian construction industry: Strengths, weaknesses, and influential demographic characteristics." *Buildings* 12: 639. https://doi.org/10.3390/buildings12050639.

Kalteh, H. O., S. B. Mortazavi, E. Mohammadi, and M. Salesi. 2021. "The relationship between safety culture and safety climate and safety performance: A systematic review." *Int. J. Occup. Saf. Ergon.* 27 (1): 206–216. https://doi.org/10.1080/10803548.2018.1556976.

Khaire, U. M., and R. Dhanalakshmi. 2022. "Stability of feature selection algorithm: A review." *J. King Saud Univ. Comput. Inf. Sci.* 34: 1060–1073. https://doi.org/10.1016/j.jksuci.2019.06.012.

Kochenderfer, M. J., and T. A. Wheeler. 2019. *Algorithms for optimization*. London: The MIT Press.

Lestari, F., R. Y. Sunindijo, M. Loosemore, Y. Kusminanti, and B. Widanarko. 2020. "A safety climate framework for improving health and safety in the Indonesian construction industry." *Int. J. Environ. Res. Public Health* 17 (20): 7462. https://doi.org/10.3390/ijerph17207462.

Li, J., K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu. 2018. "Feature selection: A data perspective." *ACM Comput. Surv.* 50 (6): 1–45. https://doi.org/10.1145/3136625.

Loosemore, M., R. Y. Sunindijo, F. Lestari, Y. Kusminanti, and B. Widanarko. 2019. "Comparing the safety climate of the Indonesian and Australian construction industries: Cultural and institutional relativity in safety research." *Eng. Constr. Archit. Manage.* 26 (10): 2206–2222. https://doi.org/10.1108/ECAM-08-2018-0340.

Norman, G. 2010. "Likert scales, levels of measurement and the "laws" of statistics." *Adv. Health Sci. Educ.* 15 (5): 625–632. https://doi.org/10.1007/s10459-010-9222-y.

Nzuva, S., and L. Nderu. 2019. "The superiority of the ensemble classification methods: A comprehensive review." *J. Inf. Eng. Appl.* 9 (5): 43–53. https://doi.org/10.7176/jiea/9-5-05.

Pedregosa, F., et al. 2011. "Scikit-learn: Machine learning in Python." *J. Mach. Learn. Res.* 12: 2825–2830.

Sagi, O., and L. Rokach. 2018. "Ensemble learning: A survey." *WIREs Data Min. Knowl. Discovery* 8 (4): e1249. https://doi.org/10.1002/widm.1249.

Schumacker, R. E., and R. G. Lomax. 2015. *A beginner's guide to structural equation modeling*. 4th ed. Abingdon: Taylor & Francis.

Solgi, R. (Mohammad). n.d. *Geneticalgorithm by R.M. Solgi*.

Solorio-Fernández, S., J. A. Carrasco-Ochoa, and J. F. Martínez-Trinidad. 2020. "A review of unsupervised feature selection methods." *Artif. Intell. Rev.* 53 (2): 907–948. https://doi.org/10.1007/s10462-019-09682-y.

Stevens, S. S. 1946. "On the theory of scales of measurement." *Science* 103 (2684): 677–680. https://doi.org/10.1126/science.103.2684.677.

Sunindijo, R. Y., M. Loosemore, F. Lestari, Y. Kusminanti, B. Widanarko, C. Febrina, and F. Amelia. 2019. "Comparing safety climate in infrastructure and building projects in Indonesia." *MATEC Web Conf.* 258: 02024. https://doi.org/10.1051/matecconf/201925802024.

Sunindijo, R. Y., and P. X. W. Zou. 2012. "Political skill for developing construction safety climate." *J. Constr. Eng. Manage.* 138 (5): 605–612. https://doi.org/10.1061/(asce)co.1943-7862.0000482.

Tharwat, A. 2021. "Classification assessment methods." *Appl. Comput. Inf.* 17 (1): 168–192. https://doi.org/10.1016/j.aci.2018.08.003.

Widjaja, T. R. A., V. S. Indarto, and R. Wulandari. 2020. "Factors that influence safety performance in Indonesia: Case study in semarang city construction industry." *Adv. Nat. Appl. Sci.* 14 (1): 154–159.

Widyanty, W., A. Daito, S. Riyanto, and D. Nusraningrum. 2020. "Gaining a competitive advantage through strategic human resource management in Indonesian construction industry." *Manage. Sci. Lett.* 10 (9): 2021–2028. https://doi.org/10.5267/j.msl.2020.2.010.

Yip, M. W., A. H. H. Ng, and D. H. C. Lau. 2012. "Employee participation: Success factor of knowledge management." *Int. J. Inf. Educ. Technol.* 2 (3): 262–264. https://doi.org/10.7763/ijiet.2012.v2.125.

Zhang, H. 2004. "The optimality of Naive Bayes." In Vol. 2 of *Proc., 17th Int. Florida Artificial Intelligence Research Society Conf.* 562–567.

Zohar, D. 1980. "Safety climate in industrial organizations: Theoretical and applied implications." *J. Appl. Psychol.* 65 (1): 96. https://doi.org/10.1037/0021-9010.65.1.96.

Zohar, D. 2010. "Thirty years of safety climate research: Reflections and future directions." *Accid. Anal. Prev.* 42 (5): 1517–1522. https://doi.org/10.1016/j.aap.2009.12.019.