APPLICATION OF MACHINE LEARNING METHODS FOR PREDICTION OF SEAFARER SAFETY PERCEPTION

Reference NO. IJME 725, DOI: 10.5750/ijme.v164iA3.725

B Arslanoğlu* G Elidolu T Uyanık, İstanbul Technical University, Turkey

*Corresponding Author. B. Arslanoğlu, demirkolb@itu.edu.tr

KEY DATES: Submitted: 30/04/21; Final acceptance: 28/11/22; Published: 31/01/23

SUMMARY

This study aims to predict seafarer safety perceptions and evaluate their feedback to understand the human factor in a ship's safety with machine learning algorithms. A questionnaire survey has been conducted with 304 seafarers' participation and they responded to several safety climate and perception indicators that are based on literature, for instance, safety assessment of supervisors and company, company's training arrangement, accident and near-miss reporting etc**.** Scores of survey results have been estimated with four machine learning algorithms, namely multiple linear regression, support vector regression, random forest and decision tree regression. According to the findings, the multiple linear regression method gave the best prediction performance for seafarer safety perception level with a 4.07 mean absolute percentage error. It was seen that the machine learning techniques can be applied in the prediction of seafarer safety perception based on collected data. This study may provide useful perspectives for maritime companies in improving safety of ships.

KEYWORDS

Machine Learning; Safety; Safety Climate; Seafarer, Human Factor.

1. INTRODUCTION

Maritime transportation is the most widely preferred mode for world trade (UNCTAD, 2019) and it involves various risks due to the working environment. Despite the increase in international rules, regulations and inspections to minimize those risks, marine accidents still are one the serious problems in this leading industry. Accidents that occur on vessels are mostly collision, grounding, cargo operation or maintenance failures and these have great potential to cause loss of lives, damage to property and the environment. According to European Maritime Safety Agency records, 19418 casualties and incidents occurred on cargo, passenger, service, fishing and other types of vessels between 2014-2019, and 493 of these are considered very serious casualties. In total, 6210 persons were injured and 496 persons lost their lives in the same period (EMSA, 2020).

The responsible authority for global maritime issues, the International Maritime Organization (IMO), has introduced many rules and regulations in response to accidents and to increase the safety of ships. Especially, the International Convention for the Safety of Life at Sea (SOLAS) is considered one of the four pillars of the industry and includes regulations for vessels to meet the minimum safety standards (IMO, 2014). In addition to the technical deficiencies of the ships, the human factor had also a considerable impact on the accidents, therefore the International Safety Management Code (ISM) also entered into force in 1998 as a part of SOLAS. The Code has brought important regulations for companies, it requires building Safety Management System for reducing human errors by implementing and developing the safety procedures on their fleet vessels. Working environments on board have become safer in line with these regulations, however, human factors still have a major role in most the marine accidents (Fan et al., 2020).

In workplaces, employees should have a reliable safety perception to reduce human errors (Lu et al., 2008). The safety perception will be shaped by occupational training and experience, as well as directly proportional to the safety level of the environment where the personnel work.

In this issue, a concept named safety climate comes forward for a better understanding of the level of workplace safety. It is the atmosphere formed as a result of general attitudes, opinions, perceptions and behaviours about safety in a working environment. This concept has been introduced by Dov Zohar for the first time in literature, and it is explained as the total perceptions and opinions of people in a working area (Zohar, 1980). Another definition is a concept in which employees believe, they could be supported and awarded by their managers through the implementation of safety-related procedures in the workplace (Hofmann, 1998). According to Neal and Griffin, safety climate is a personal notion that is affected by workplace-related procedures and policies which are obeyed for ensuring safety (Neal and Griffin, 1998). Although the studies on this concept have mostly been carried out in various industries such as manufacturing, construction, shipbuilding etc. (Kim et al., 2017; Oah et al., 2018; Patel et al., 2015), it can also be applied in the maritime field (Lu et al., 2008).

Ships are riskier working environments than land workplaces, due to their structural characteristics and dynamic atmospheric conditional areas they operate (Oldenburg et al. 2010). This situation still causes many occupational accidents or major marine accidents despite various rules and regulations. These incidents may be eliminated by observing the safety climate onboard, therefore, this paper aims to predict seafarers' safety perceptions. In this context, a questionnaire was prepared based on safety climate determinants that were used in literature by adapting each item to seafarers. The survey results were evaluated and the safety perception of seafarers was predicted. In this stage, machine learning is applied because it is a strong method in the artificial intelligence (AI) field that has been proven to be effective for many applications (Uyanık et al., 2019). The systems supported with AI methods used in various fields comprising image recognition, surveillance, healthcare, fraud prevention, tourism marketing, trading and shipping etc. help in both saving time and obtaining reliable data (Alsheikh et al., 2014). On the other hand, there are not sufficient studies using these methods in the maritime industry. Most of the papers are concerned with efficiency problems on vessels, for instance, navigation, route prediction, fuel consumption (Abebe et al., 2020; Kim et al., 2017; Uyanık et al., 2020). There is a gap in safety-related issues utilizing machine learning techniques according to the authors' knowledge. Therefore, studies using these methods in different industries have been reviewed mostly.

Davoudi et al. used machine learning in the agribusiness sector. They collected data from an insurance company which were more than 33,000 workers' job-related accidents and tested the performance of machine learning techniques in predicting the severity of accidents' consequences (Davoudi et al., 2019). Rawson & Brito used machine learning methods in the field of maritime risk assessment (Rawson and Brito, 2022). Veerappa et al., investigated vessel type classification with an explainable AI method (Veerappa et al. 2022). Luo et al., applied machine learning methods to ship wake prediction for shoreside protection (Luo et al., 2022). Uyanik et al., investigated the application of machine learning techniques for visibility prediction to safer navigation (Uyanık et al., 2021). Sarkar et al., conducted a study that developed a model with machine learning techniques for the estimation of occupational accidents in a steel plant. With the application of support vector machine and artificial neural network techniques and based on occupational accidental data, they estimated injuries, near-misses and property damages (Sarkar et al., 2019). Another research is about developing a model with an artificial neural network (ANN) for estimation of the workers' safe work behaviour. Patel and Jha first applied a questionnaire to construction site workers in India. In the questionnaire, there were 10 safety climate determinants that were used in other related studies in the literature. These determinants were used as inputs for the ANN model and they predicted employees' safe work behaviours as outputs (Patel and Jha, 2015). In Korea, Kang and Ryu developed a machine learning-based model to minimize and prevent occupational accidents which happen on construction sites. In this model, they used the random forest technique to predict the types of accidents and the model showed a 71.3% estimation performance (Kang and Ryu, 2019).

In this study, it was aimed to contribute to the literature and provide a different perspective for maritime/crew companies to understand the safety climate in their vessels and evaluate the crew performance accordingly. The remainder of the paper is organized as follows. In section 2, data collection and the machine learning techniques are explained. Section 3 presents the prediction performance with the success levels. Section 4 discusses the research results, limitations and further study aspects. Finally, Section 5 concludes the paper.

2. MATERIAL AND METHOD

Firstly, the questionnaire was prepared for understanding the safety perceptions of seafarers. It was sent to the seafarers and responses were collected over a period of time. Data including their personal information and safety levels from the survey were organized to be fit for the system. In this stage, the dataset was divided into two subsets as "training" and "test" data. Four techniques that were used for prediction were trained, following this phase, the remainder of data was tested and compared with the actual safety perception levels. Using error metrics, the rate of estimation success of each technique was assessed. Finally, the techniques were compared with each other and was evaluated the most suitable technique for the prediction. The steps of the methodology for the seafarer safety perceptions prediction model are illustrated in Figure 1 for a better understanding of the study.

2.1 SURVEY PARTICIPANTS AND DATA COLLECTION

In this research, a survey has been conducted for data collection. It was sent to Turkish seafarers who have oceangoing experience between 1 year and 20 years. Ranks are defined as master, chief engineer, chief officer, deck officer, engineer, cadet/student, ratings including able seaman, oiler, electrician, fitter, and pumpman. They were asked to state their age, marital status, gender, rank, sea experience, and educational degree. The majority of the respondents are male with a ratio of 86%. Data collection was carried out from April 2019 to February 2020. The survey has been distributed to a total of 450 seafarers, but it could reach 304 seafarers' data due to vessels' intense work conditions and communication challenges. Therefore, the proper response rate is 67,55% level at the end of the survey. To prevent reporting biases as much as possible, it was collected responses from the seafarers anonymously.

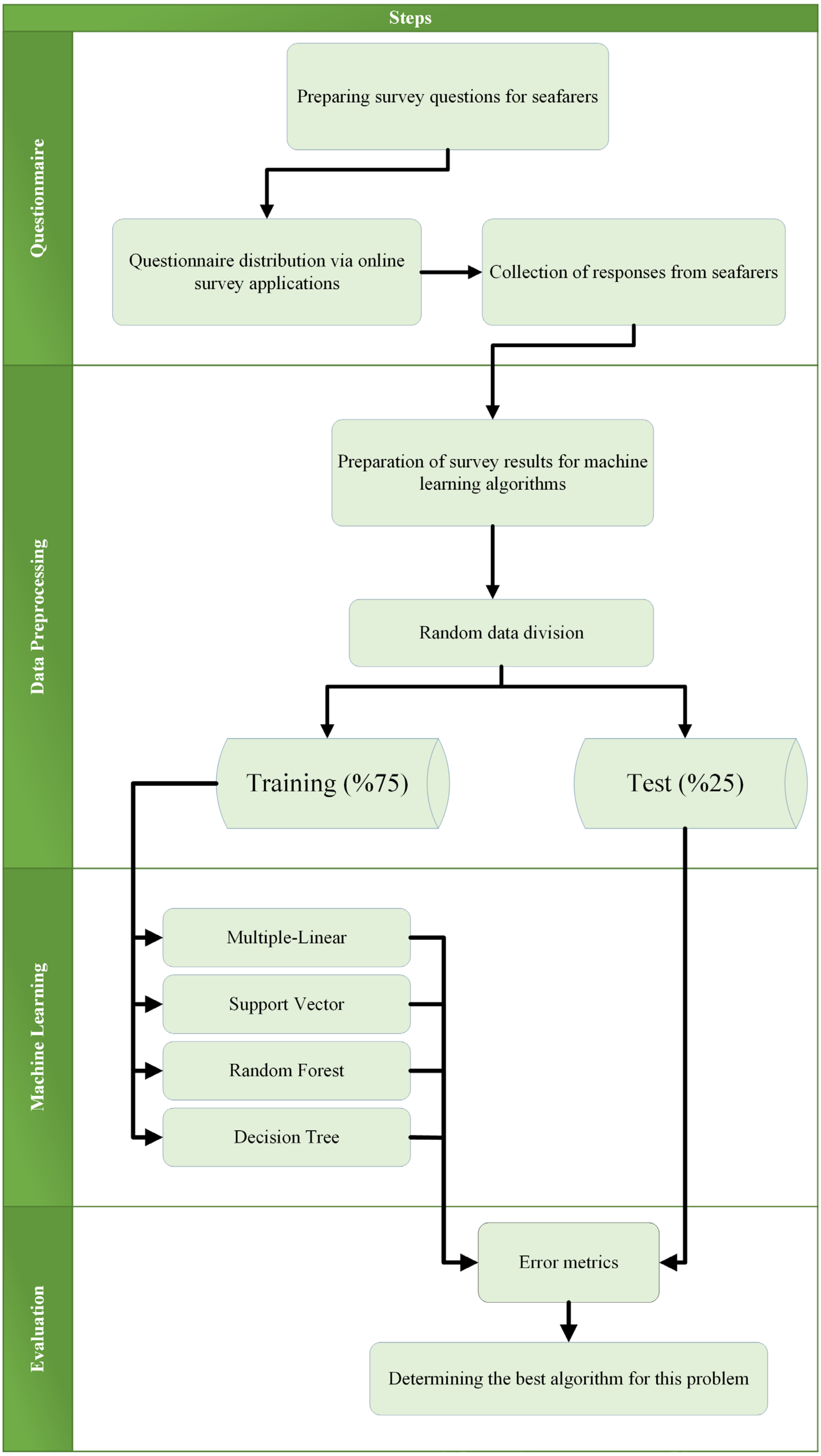

Figure 1. A prediction model for seafarer safety perceptions.

The context of the questionnaire was prepared based on the studies analysing the safety climate and safety perceptions in the literature (Glendon and Litherland, 2001; Lu and Tsai, 2010; Mearns et al., 2003; Williamson et al., 1997). The questions are in nine main parts which are ship supervisors' safety assessment, company's safety assessment, company's training planning, co-worker interaction, workload and pressure, reporting of near misses and occupational accidents, straightforwardness of safety rules and procedures, safety assessment of the job, and employees' safety perception and fatalism. A total of 51 questions were asked on a five-point Likert scale (1=strongly disagree, 2=disagree, 3=nor agree or disagree, 4=agree, 5=strongly agree). Also, Cronbach's alpha was calculated to assess the internal consistency of the scales used in the study (Taber, 2018). As seen in Table 1, Cronbach's alpha values vary between 0.745 and 0.941 and this means the survey could be considered internal consistent, since the value should be 0.70 and above according to literature (Lu and Tsai, 2010).

2.2 SURVEY SECTIONS

The safety climate and safety perception indicators used in the survey were illustrated in Table 1. It includes Cronbach's alpha values, the number of items and a sample item for each part.

The system outputted the safety perception levels with the application of four different machine learning methods, which are frequently used and generally yield successful results (Kushwah et al., 2021; Rosa et al., 2019; Wang et al., 2022; Zhang et al., 2020; Jiang et al., 2021). Table 2 illustrates the variables that are used in machine learning applications.

2.3 MACHINE LEARNING METHODS

Machine learning algorithms have been functional and effective in classification, prediction, recognition and identification problems in recent years. In this work,

Table 1: Survey sections

Part name	Sample item	No.	CA
Safety Assessment of Ship Supervisors	My supervisors care about the ship crew's safety.	8	0.898
Safety Assessment of Company	My company has a regular job safety meeting.	10	0.884
Company's Training Arrangement	The safety training programmes in my company help prevent accidents.	6	0.932
Workload and Work Pressure	There are enough seafarers on board to carry out the required work.	2	0.763
Reporting of Occupational Accidents	The seafarers are willing to report near misses.	5	0.899
Safety Assessment of Co-workers	My coworkers care about others' safety.	5	0.941
Straightforwardness of Safety Rules and Procedures	Rules about safety are not difficult to understand.	3	0.832
Safety Assessment of Job	One who works onboard can easily get hurt.	4	0.745
Safety Perception and Fatalism	I follow safety rules even under intense work pressure.	8	0.839

Table 2: Sample data set including variables.

Position	Personal Data					
	Experience	Age	Gender	Education	Marital Status	Safety Score
Engine Officer	3	26	M	3	S	4,142
Able seamen	4	28	M	1	M	4,428
Chief Engineer	10	33	M	3	M	3,690
Cadet/Student	1	30	F	3	S	3,809
1st assist Eng.	5	28	M	3	S	4,690
Deck Officer	2	26	F	4	M	3,357
2nd Engineer	5	29	M	3	M	3,547
Chief Officer	10	39	M	3	M	4,470
Master	19	43	M	3	M	4,595
Electrician	20	37	M	1	M	3,023

seafarers' safety perceptions were estimated by machine learning techniques. Data provided as input to the system are the seafarers' position onboard, sea experience, age, gender, educational degree, marital status and survey scores.

2.3(a) Multiple-Linear Regression

Multiple or multiple-linear regression (MLR) is used in the stage of prediction with two or more independent or predictor variables.

$$y = a+b_1x_1+b_2x_2+\cdots+b_nx_n \tag{1}$$

In Equation (1), y is the dependent variable, where a is the y-axis intercept of the regression curve, bi is the coefficient of the estimate value x_i, for $1 \leq i \leq n$.

2.3(b) Support Vector Regression

Support vector machine is one of the supervised learning models for classification and regression [34]. Support vector regression uses two linear or non-linear vectors around the real data for forecasting. The kernel functions are used for regression which resembles support vector machines but SVR adjusts the permissible limits (ε) (Harris et al., 1997).

In Figure 2, the hyperplane represents the real values. Yellow dots show support vector machines. The cyan dots symbolize the predictions that are in the threshold range (+Ɛ/− Ɛ) when the green dots are used for being out of range.

2.3(c) Decision Tree Algorithm

Decision tree (DT) learning, a method for approaching separate-valued goal functions, is frequently used in machine learning studies. In this method, samples are symbolized as attribute-value pairs. The objective function of the decision tree has separate output values that designate a category for each example and properly appoint new input. Decision tree may be useful while executing the selection of automatic property and lessening complications. The decision tree method (DTs) solves regression and classification problems effectively. One of the important characteristics of DTs is their low computational complexity. Additionally, no suppositions are needed for the parameters' dispersion of the predictor. The DTs approach is also powerful in using insufficient data (Quinlian, 1986; Rokach and Maimon, 2005). A simple demonstration of the technique is in Figure 3.

2.3(d) Random Forest Algorithm

Random Forest (RF) is a trending machine learning algorithm mostly employed in classification and regression implementations as a practical tool. Figure 3 demonstrates the technique simply. In this study, the proposed random forest algorithm is like bagged DTs.

A random forest is a sum of classifiers that are used as tree-based. $h(X;\Theta_k),k=1,2,..,K$, as independent variables, Θ_k are similarly spread out random vectors and X is the input and Y is the output, so that (X, Y) shapes the real data. Firstly, K samples are received from the real training data (X, Y) with the process of bootstrap, each sample has an identical format as that in the original data. Then, K regression patterns are constructed of the decision tree for all samples to get K regression estimation outcomes. At the end of the process, the mean of K results is the last prediction result.

RF algorithm improves gaps in the regression models by constituting dissimilar training data to support the extrapolation estimation capability of the associated

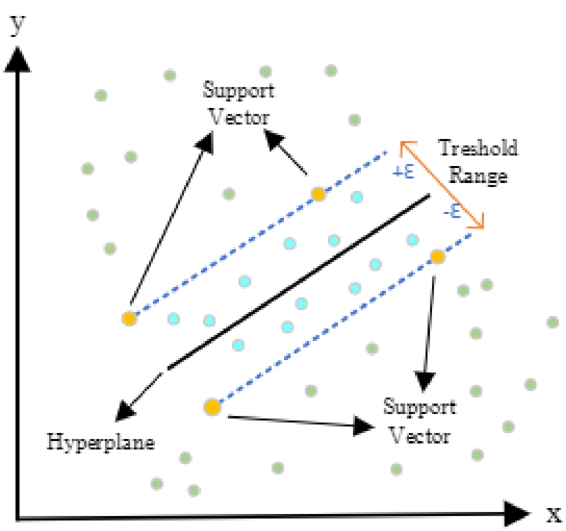

Figure 2. Support vector machine

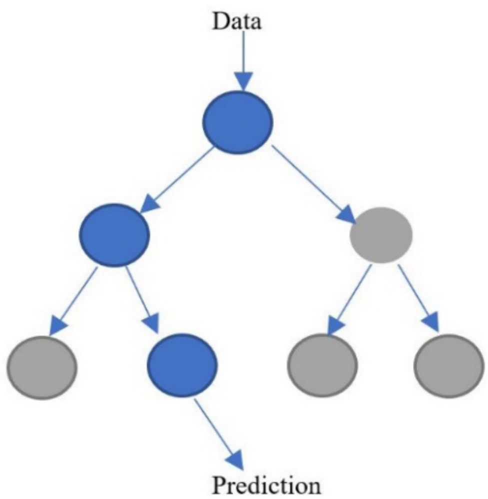

Figure 3. Decision tree algorithm

regressions. Following K's training, the regression prediction model may be obtained from series $h(X,\theta_1), h(X,\theta_2), \ldots, h(X,\theta_k)$ that are created from a combined forecasted pattern. The last prediction result is expressed in Equation 2 (Fernandez-Gonzalez et al., 2019; Wehenkel et al., 2006).

$$h(x) = \left(\frac{1}{K}\right)\sum_{k=1}^{K} h(X;\theta_k) \tag{2}$$

2.3(e) Validation of the ML Techniques

In literature, the K-fold cross-validation method is used to assess the success of prediction methods and also prevent overfitting (Lazakis et al., 2019). The basic rationale in the application of this method is to divide the data set into k equal parts and use each part as both training and test data in all iterations. In this study, the ML techniques are validated via this method with the number of k iterations is adjusted as 4, as illustrated in Figure 5.

2.4 ERROR METRICS

It was used different types of error metrics to evaluate the success of the introduced algorithms on the problem. These are root mean squared error (RMSE), mean absolute error (MAE) and mean percentage error.

2.4(a) Root Mean Squared Error

The RMSE is expressed as in Eq. (3).

$$RMSE = \sqrt{\frac{\sum_{j=1}^{n} e_j^2}{n}} \tag{3}$$

where, $e_j = k_{(j)actual} - k_{(j)predicted}$ and, n is size of the data set, $k_{(j)actual}$ is the real value, $k_{(j)predicted}$ is the predicted value.

2.4(b) Mean Absolute Error

The MAE is the average of the absolute value of errors that was given in Eq. (4).

$$MAE = \frac{1}{n}\sum_{j=1}^{n} \left|e_j\right| \tag{4}$$

where n is the size of the data set, $e_j = k_{(j)actual} - k_{(j)predicted}$, $k_{(j)actual}$ is real value, $k_{(j)predicted}$ is predicted value.

2.4(c) Mean Absolute Percentage Error

The main formula for the MAPE can be found in Eq. (5).

$$MAPE = \sum_{j=1}^{n} \frac{e_j}{k(j)_{actual}} \tag{5}$$

Where n is size of the data set, $e_j = k_{(j)actual} - k_{(j)predicted}$, $k_{(j)actual}$ is real value, $k_{(j)predicted}$ is predicted value.

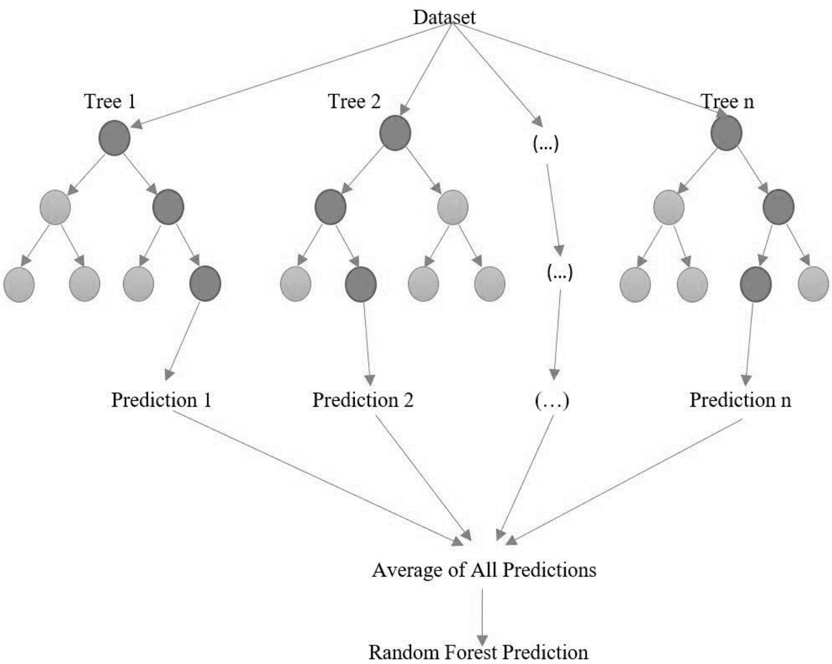

Figure 4. Random forest algorithm

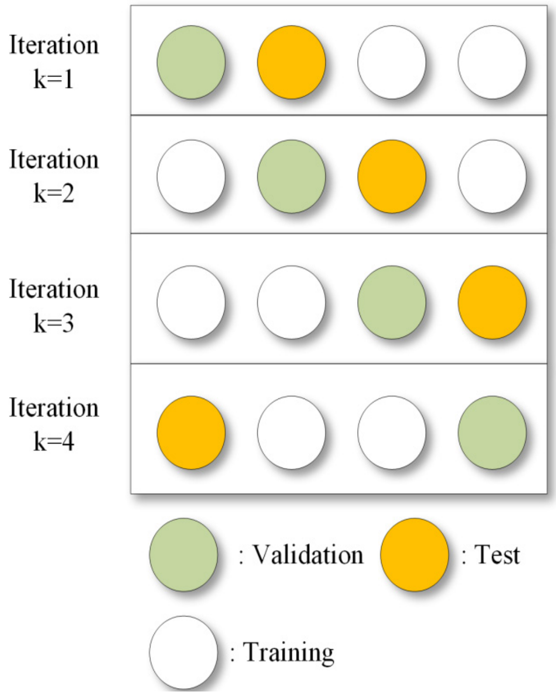

Figure 5. K-fold cross-validation for 4 data packs and k=4.

3. FINDINGS AND RESULTS

In this paper, the data size is made up of 304 participants. 228 participants of this data were selected randomly for training data and taught to a computer. The remaining 76 participants were used as test data and asked to be predicted by the computer. It was used the Python programming language version 3.0 in the application of the case study section. Codes for the study were written in Spyder 3.0 interface. Pandas (Kapadia et al., 2019) and Numpy (McClaren, 2018), Sklearn (scikit-learn) (Erickson et al., 2018), libraries of Python programming language are also used for the case study. Figure 6 illustrates the simulation results of randomly selected 50 samples with four different machine learning methods for safety perception prediction.

In Figure 6, the points indicated by the red dots symbolize the actual safety perception values according to the survey results which were used as real data input. In this step, all scores were valued between 1 and 5 for each seafarer. The other colours represent the estimation values achieved by prediction techniques. It was applied the classical and the most widely used techniques and aimed to find the method that performs the highest success on the data set.

Table 3 illustrates error rate comparisons of the machine learning techniques used to predict safety perception. Considering all methods, multiple linear regression is more advantageous than other techniques for safety perception prediction. Its success rate is 95.92%.

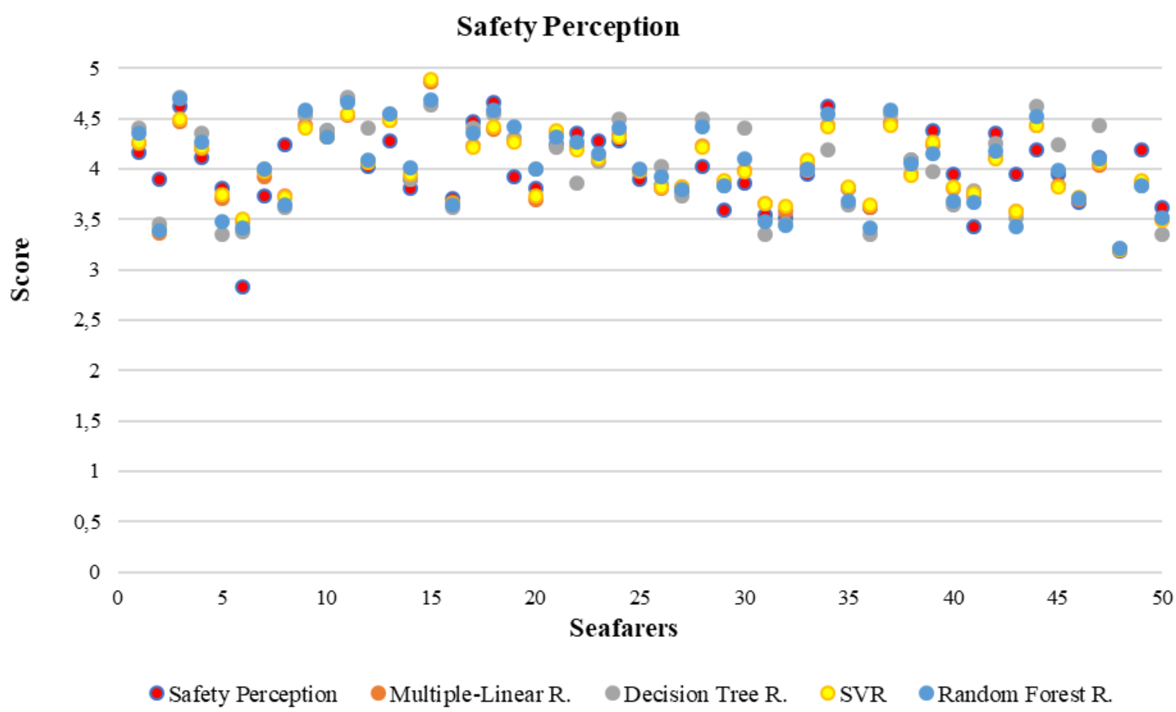

Figure 6. Comparison of four machine learning methods for prediction of seafarers' safety perception

Table 3: Error rates for estimation of safety perception.

Regression Method	Metrics		
	RMSE	MAE	%MAPE
Decision Tree	0,219167	0,084345	6,014093
Multiple Linear	0,159656	0,044643	4,071552
Random Forest	0,183206	0,056815	4,686330
Support Vector	0,163374	0,071347	4,177861

Table 4: Validation scores for estimation of safety perception.

Regression Method	Validation Score (RMSE)				
	Iter.1	Iter.2	Iter.3	Iter.4	Mean
Decision Tree	0,25896	0,24731	0,23983	0,25109	0,24930
Multiple Linear	0,18705	0,17986	0,19250	0,17731	0,18418
Random Forest	0,21654	0,21367	0,20998	0,21536	0,21389
Support Vector	0,18750	0,19856	0,17999	0,18447	0,18763

K-fold cross-validation was made for all prediction techniques. The data set was divided into 4 subsets accordingly. All techniques were validated with the calculation of RMSE values. Table 4 illustrates the validation scores.

Although decision tree regression performed less than the other three techniques, it achieved a 93.98% success rate in safety perception level estimation. As a result of the research, it was found that the use of multiple linear regression and support vector regression techniques could yield more reliable prediction results. The simulation results have shown that developed machine learning models in this paper are useful for seafarers' safety perception prediction.

4. DISCUSSION

In this paper, it was aimed to estimate the safety perceptions of seafarers working on Turkish flagged vessels. Since the machine learning methods could make more accurate predictions than the classic statistical methods in proper circumstances, testing this ability may be interesting in the maritime industry. To obtain data, was applied a questionnaire to seafarers to respond to safety-related issues. Although it has been used in many studies in the literature, this survey method that measures the safety perceptions of employees may not be sufficient, so this could be the limitation of the paper. A more comprehensive examination of vessels should be conducted with experts to analyse crew safety climate more effectively. It is difficult for many companies to allow onboard research since operating their ships with extra people for a long time is often costly and arduous. Therefore, in the face of these challenges, it was preferred to obtain data with the questionnaire method. Another limitation is the difficulty of contact with seafarers. It is not an easy process to reach seafarers actively working onboard, due to the intense working conditions and communication problems. These problems prevented to increase in the number of samples. For further study, the authors will continue to research to increase and improve the input data that is used in the machine learning process in addition to the questionnaire survey. Also, increasing the number of data by collecting more samples from seafarers will help the applied techniques to be trained with more data, thereby increasing the prediction performance.

5. CONCLUSION

Shipping has great importance in the matter of world trade; therefore, the seafaring profession is essential to execute this economic activity. With the contributions of many studies analysing ship accidents' causes, it has been learned how effective the human factor is in the occurrence of accidents. The consequences of ship-related accidents can lead to enormous disasters, both in terms of human life and the natural environment. To prevent ship-related or occupational accidents caused by human error, the safety perception of the ship's personnel should be well understood. Based on the employee profile examination by using machine learning-supported systems, it may be easier that the relevant authorities develop new training programmes and eliminate the deficiencies of their employees, thus ensuring the safer operation of ships.

This work aims to make contributions to improve safety on ships, with the application of machine learning method which was not used in other related studies. By introducing the data obtained from the questionnaire survey to a machine learning-assisted system, seafarers' safety perception levels have been predicted successfully. It should be remembered that the performance and success of the predictions will advance with an increasing number of ship personnel data included in the system.

6. REFERENCES

1. ABEBE, M., SHIN, Y., NOH, Y., LEE, S., & LEE, I. (2020). *Machine learning approaches for ship speed prediction towards energy efficient shipping.* Applied Sciences (Switzerland), 10 (7). https://doi.org/10.3390/app10072325
2. ALSHEIKH, M. A., LIN, S., NIYATO, D., & TAN, H. P. (2014). *Machine learning in wireless sensor networks: Algorithms, strategies, and applications.* IEEE Communications Surveys and Tutorials, 16(4), 1996–2018. https://doi.org/10.1109/COMST.2014.2320099
3. DAVOUDI K., F., FREEMAN, S. A., & MOSHER, G. A. (2019). *Evaluating machine learning performance in predicting injury severity in agribusiness industries.* Safety Science, 117, 257–262. https://doi.org/10.1016/j.ssci.2019.04.026
4. EMSA. (2020). Accident Investigation - Annual Overview - EMSA - European Maritime Safety Agency. *Annual Overview of Marine Casualties and Incidents 2020.* Accessed 01.03.2021. http://www.emsa.europa.eu/accident-investigation-publications/annual-overview.html
5. ERICKSON, C. B., ANKENMAN, B. E., & SANCHEZ, S. M. (2018). *Comparison of Gaussian process modelling software.* European Journal of Operational Research, 266(1), 179–192. https://doi.org/10.1016/j.ejor.2017.10.002
6. FAN, S., ZHANG, J., BLANCO-DAVIS, E., YANG, Z., & YAN, X. (2020). *Maritime accident prevention strategy formulation from a human factor perspective using Bayesian Networks and TOPSIS.* Ocean Engineering, 210(June), 107544. https://doi.org/10.1016/j.oceaneng.2020.107544
7. FERNANDEZ-GONZALEZ, P., BIELZA, C., & LARRANAGA, P. (2019). *Random Forests for Regression as a Weighted Sum of k-Potential Nearest Neighbours.* IEEE Access, 7, 25660–25672. https://doi.org/10.1109/ACCESS.2019.2900755
8. GLENDON, A. I., & LITHERLAND, D. K. (2001). *Safety climate factors, group differences and safety behaviour in road construction.* Safety Science, 39(3), 157–188. https://doi.org/10.1016/S0925-7535(01)00006-6
9. HARRIS D., BURGES, C. C., KAUFMAN, L., SMOLA, A.J., VAPNIK, V. N. (1997); *"Support Vector Regression Machines".* Advances in Neural Information Processing Systems 9, NIPS 1996, 155–161, MIT Press.
10. HOFMANN, D. A., & STETZER, A. (1998). *The Role Of Safety Climate And Communication In Accident Interpretation: Implications For Learning From Negative Events.* Academy of Management Journal, 41(6), 644–657. https://doi.org/10.2307/256962
11. IMO, (2014). International Maritime Organisation *SOLAS, consolidated edition 2014: Consolidated text of the International Convention for the Safety of Life at Sea, 1974, and its Protocol of 1988*: articles, annexes and certificates.
12. JIANG Y., XUE-GANG W., ZAO-JIAN Z., ZHAO-LONG Y., (2021), *Identification of coupled response models for ship steering and roll motion using support vector machines,* Applied Ocean Research, 110, 102607, 0141-1187, https://doi.org/10.1016/j.apor.2021.102607.
13. KANG, K., & RYU, H. (2019). *Predicting types of occupational accidents at construction sites in Korea using random forest model.* Safety Science, 120, 226–236. https://doi.org/10.1016/j.ssci.2019.06.034
14. KAPADIA, K., ABDEL-JABER, H., THABTAH, F., & HADI, W. (2019). *Applied Computing and Informatics Sport analytics for cricket game results using machine learning: An experimental study.* Applied Computing and Informatics. https://doi.org/10.1016/j.aci.2019.11.006
15. KIM, K. W., PARK, S. J., LIM, H. S., & CHO, H. H. (2017). *Safety Climate and Occupational Stress According to Occupational Accidents Experience and Employment Type in Shipbuilding Industry of Korea.* Safety and Health at Work, 8(3), 290–295. https://doi.org/10.1016/j.shaw.2017.08.00
16. KIM, J. S., & JEONG, J. S. (2017). *Extraction of reference seaway through machine learning of ship navigational data and trajectory.* International Journal of Fuzzy Logic and Intelligent Systems, 17(2), 82–90. https://doi.org/10.5391/IJFIS.2017.17.2.82
17. KUSHWAH J.S., ATUL K., SUBHASH P., RISHI S., AMOL G., SHYAM G., (2021), *Comparative study of regressor and classifier with decision tree using modern tools*. Materials Today: Proceedings, 2214-7853, https://doi.org/10.1016/j.matpr.2021.11.635.
18. LAZAKIS, I., GKEREKOS, C., & THEOTOKATOS, G. (2019). *Investigating an SVM-driven, one-class approach to estimating ship systems condition.* Ships and Offshore Structures, 14(5), 432–441. https://doi.org/10.1080/17445302.2018.1500189
19. LU, C.-S., & TSAI, C.-L. (2008). *The effects of safety climate on vessel accidents in the container shipping context.* Accident Analysis & Prevention, 40(2), 594–601. https://doi.org/10.1016/J.AAP.2007.08.015
20. LU, C. S., & TSAI, C. L. (2010). *The effect of safety climate on seafarers' safety behaviours in container shipping.* Accident Analysis and Prevention, 42(6), 1999–2006. https://doi.org/10.1016/j.aap.2010.06.008
21. LUO Y., CHENG Z., JUNLIANG L., HUANLIN X., FENGHUA Z., DONGXIAO W., XIAOMIN

L., SHENGAN W., WEIQIANG W., FENGYAN S.(2022). *Identifying ship-wakes in a shallow estuary using machine learning,* Ocean Engineering, 246, 110456, 0029-8018, https://doi.org/10.1016/j.oceaneng.2021.110456.

22. MCCLARREN, R. (2018). *Computational Nuclear Engineering and Radiological Science UsingPython.* https://doi.org/10.1016/B978-0-12-812253-2.00005-4
23. MEARNS, K., WHITAKER, S. M., & FLIN, R. (2003). *Safety climate, safety management practice and safety performance in offshore environments.* Safety Science, 41(8), 641–680. https://doi.org/10.1016/S0925-7535(02)00011-5
24. NEAL, A., & GRIFFIN, M. A. (2006). *A study of the lagged relationships among safety climate, safety motivation, safety behaviour, and accidents at the individual and group levels.* Journal of Applied Psychology, 91(4), 946–953. https://doi.org/10.1037/0021-9010.91.4.94
25. OAH, S., NA, R., & MOON, K. (2018). *The Influence of Safety Climate, Safety Leadership, Workload, and Accident Experiences on Risk Perception: A Study of Korean Manufacturing Workers.* Safety and Health at Work, 9(4), 427–433. https://doi.org/10.1016/j.shaw.2018.01.008
26. OLDENBURG, M., BAUR, X., & SCHLAICH, C. (2010). *Occupational risks and challenges of seafaring.* Journal of Occupational Health, 52(5), 249–256. https://doi.org/10.1539/joh.k10004
27. PATEL, D.A., & JHA, K.N. (2015). *Neural Network Approach for Safety Climate Prediction.* Journal of Management in Engineering, 31(6), 05014027. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000348
28. QUINLAN, J. R. (1986). *Induction of decision trees.* Machine Learning, 1(1), 81–106. https://doi.org/10.1007/bf00116251
29. RAWSON A., BRITO M. (2022). *A survey of the opportunities and challenges of supervised machine learning in maritime risk analysis,* Transport Reviews, 0144-1647, https://doi.org/10.1080/01441647.2022.2036864.
30. ROKACH, L., & MAIMON, O. (2005). *Top-down induction of decision trees classifiers - A survey.* IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, 35(4), 476–487. https://doi.org/10.1109/TSMCC.2004.843247
31. ROSA M. A.C., IVAN C. R., CHRISTOPHER E., BELEN M.B., JOSE M. M.V., (2019). *Using a decision tree algorithm to predict the robustness of a transshipment schedule,* Procedia Computer Science, 149, 529-536, 1877-0509, https://doi.org/10.1016/j.procs.2019.01.172.
32. TABER, K.S. (2018). *The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education*. Res Sci Educ 48, 1273–1296. https://doi.org/10.1007/s11165-016-9602-2
33. SARKAR, S., VINAY, S., RAJ, R., MAITI, J., & MITRA, P. (2019). *Application of optimized machine learning techniques for prediction of occupational accidents.* Computers and Operations Research, 106, 210–224. https://doi.org/10.1016/j.cor.2018.02.021
34. SMOLA, A. J., & SCHÖLKOPF, B. (2004). *A tutorial on support vector regression.* In Statistics and Computing Vol. 14, Issue 3, pp. 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88
35. UNCTAD (2019). *Review of Maritime Transport 2019* (United Nations publication. Sales No. E.19.II.D.20. New York and Geneva).
36. UYANIK, T., ARSLANOĞLU, Y., & KALENDERLI, Ö. (2019). *Ship Fuel Consumption Prediction with Machine Learning.* 4th International Mediterranean Science and Engineering Congress. 25-27 April, Antalya, Turkey.
37. UYANIK, T., KARATUĞ, Ç., & ARSLANOĞLU, Y. (2020). *Machine learning approach to ship fuel consumption: A case of container vessel.* Transportation Research Part D: Transport and Environment, 84(May). https://doi.org/10.1016/j.trd.2020.102389
38. UYANIK, T. KARATUĞ, Ç. & ARSLANOĞLU, Y. (2021) *Machine learning based visibility estimation to ensure safer navigation in strait of Istanbul,* Applied Ocean Research, 112, 102693, 0141-1187, https://doi.org/10.1016/j.apor.2021.102693.
39. VEERAPPA M., MATHIAS A., NADIA B., MARCO F. H. (2022), *Validation of XAI explanations for multivariate time series classification in the maritime domain,* Journal of Computational Science, 58, 101539, 1877-7503, https://doi.org/10.1016/j.jocs.2021.101539.
40. WANG J., YUHAN G., YIYANG W., (2022), *A sequential random forest for short-term vessel speed prediction.* Ocean Engineering, 248, 110691, 0029-8018, https://doi.org/10.1016/j.oceaneng.2022.110691.
41. WEHENKEL, L., ERNST, D., & GEURTS, P. (2006). *Ensembles of extremely randomized trees and some generic applications.* Proc. Robust Methods for Power System State Estimation and Load Forecasting, RTE, France,1–10. https://doi.org/10.1.1.70.3238&rep=rep1&type=pdf
42. WILLIAMSON, A. M., FEYER, A. M., CAIRNS, D., & BIANCOTTI, D. (1997). *The development of a measure of safety climate: The role of safety perceptions and attitudes.* Safety

Science, 25(1–3), 15–27. https://doi.org/10.1016/S0925-7535(97)00020-9

43. ZHANG C., JUNCHI B., WELLS W., XIANG P., RUI W., RICHARD H., ZHENG L., (2020), *AIS data driven general vessel destination prediction: A random forest based approach.* Transportation Research Part C: Emerging Technologies, 118, 102729, 0968-090X, https://doi.org/10.1016/j.trc.2020.102729.

44. ZOHAR, D. (1980). *Safety climate in industrial organizations: Theoretical and applied implications.* Journal of Applied Psychology, 65(1), 96–102. https://doi.org/10.1037/0021–9010.65.1.96.

APPENDIX

Survey

		Strongly disagree	Disagree	Nor agree or disagree	Agree	Strongly agree
Safety Assessment of Ship Supervisors						
1	My supervisors care about the ship crew's safety.	1	2	3	4	5
2	My supervisors encourage safe behaviour.	1	2	3	4	5
3	My supervisors informs crew about safety rules.	1	2	3	4	5
4	My supervisors are willing to listen crew in setting the safety goals.	1	2	3	4	5
5	My supervisors follow the safety rules.	1	2	3	4	5
6	My supervisors commend safe work behaviour.	1	2	3	4	5
7	My supervisors considers the warnings and suggestions of crew about safety.	1	2	3	4	5
8	When I have a health problem, my supervisors take care of my situation and report it to the company.	1	2	3	4	5
Safety Assessment of Company						
1	My company shares safety rules and procedures with the with the seafarers.	1	2	3	4	5
2	My company shares the warnings and notifications about dangerous cargoes with the seafarers.	1	2	3	4	5
3	My company responds quickly to safety concerns and provides solutions.	1	2	3	4	5
4	My company has a regular job safety meeting.	1	2	3	4	5
5	My company's safety-related procedures are sufficient.	1	2	3	4	5
6	My company's safety procedures are reasonable.	1	2	3	4	5
7	My company supplies ships with sufficient safety equipment.	1	2	3	4	5
8	My company conducts safety inspections on ships at regular intervals.	1	2	3	4	5
9	My company rewards the crew who works safely.	1	2	3	4	5
10	My company sends me to the hospital for treatment when I have a health problem and signs me off the ship when necessary.	1	2	3	4	5
Company's Training Arrangement						
1	My company conducts an adequate number of safety training programs.	1	2	3	4	5
2	The safety training programs in my company are explicit.	1	2	3	4	5
3	The safety training programs in my company are useful.	1	2	3	4	5
4	The safety training programs in my company are worth my time.	1	2	3	4	5
5	The safety training programs in my company are compatible with my job onboard.	1	2	3	4	5
6	The safety training programmes in my company help prevent accidents.	1	2	3	4	5

Workload and Work Pressure						
1	Ship crew are given sufficient time to perform a job safely.	1	2	3	4	5
2	There are enough seafarers on board to carry out the required work.	1	2	3	4	5
Reporting of Occupational Accidents						
1	The seafarers are willing to report near misses.	1	2	3	4	5
2	The seafarers are willing to report accidents.	1	2	3	4	5
3	I do not hesitate to report a near miss or accident to my supervisors.	1	2	3	4	5
4	I learn from near misses and accidents.	1	2	3	4	5
5	My company shares the events (near misses, accidents, bad practices etc.) that happened on other ships with my ship.	1	2	3	4	5
Safety Assessment of Co-workers						
1	My co-workers care about working safely.	1	2	3	4	5
2	My co-workers follow the safety rules.	1	2	3	4	5
3	My co-workers give efforts to keep our working area safe.	1	2	3	4	5
4	My co-workers care about others' safety.	1	2	3	4	5
5	My co-workers encourage others to be safe.	1	2	3	4	5
Straightforwardness of Safety Rules and Procedures						
1	I have no difficulty understanding the purpose of the safety rules.	1	2	3	4	5
2	Rules about safety are not difficult to understand.	1	2	3	4	5
3	I know the meaning and the purpose of the International Safety Management (ISM) Code.	1	2	3	4	5
Safety Assessment of Job						
1	Working onboard involves risks.	1	2	3	4	5
2	One who works onboard can easily get hurt.	1	2	3	4	5
3	Working onboard harms crew's health.	1	2	3	4	5
4	Working onboard is unsafe.	1	2	3	4	5
Safety Perception and Fatalism						
1	Using personal protective equipment helps in preventing accidents.	1	2	3	4	5
2	I do not think the accidents are due to bad luck.	1	2	3	4	5
3	Safe operational rules and procedures can reduce accidents.	1	2	3	4	5
4	The use of safety equipment can reduce injuries and accidents.	1	2	3	4	5
5	I do not ignore safety rules to finish my work quickly.	1	2	3	4	5
6	I follow safety rules even under intense work pressure.	1	2	3	4	5
7	I do not ignore safe working procedures for convenience and comfort.	1	2	3	4	5
8	It's okay for me to get safety suggestions from others.	1	2	3	4	5