INTERNATIONAL JOURNAL OF SUSTAINABLE CONSTRUCTION ENGINEERING AND TECHNOLOGY
ISSN: 2180-3242 e-ISSN: 2600-7959

IJSCET

Vol. 15 No. 4 (2024) 53-63
https://publisher.uthm.edu.my/ojs/index.php/ijscet

Predictive Model for Incident Severity at Railway Construction Site Using Rapid Miner

Zuraidah Ngadiron[1], Reventheren Ganasan[2], Mimi Faisyalini Ramli[1]*, Mohd Eizzuddin Mahyeddin[2], M. Izzad Luqman[2], Guo Jiafu[3], N.A. Kamaluddin[4]

[1] *Power Integration System (PI Sys), Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, MALAYSIA*
[2] *Department of Transportation Engineering Technology, Faculty of Engineering Technology Universiti Tun Hussein Onn Malaysia, Pagoh, 84600, Johor, MALAYSIA*
[3] *Engineering Technology Centre of Distributed Power Generation and Smart Microgrid, Faculty of Electrical Power Engineering, Guangxi Electrical Polytechnic Institute, 530200 Yong Ning District, Nan Ning, CHINA*
[4] *School of Civil Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, MALAYSIA*

*Corresponding Author: faisyalini@uthm.edu.my
DOI: https://doi.org/10.30880/ijscet.2025.15.04.005

Article Info

Received: 21 May 2024
Accepted: 24 December 2024
Available online: 28 December 2024

Keywords

Machine learning, construction, railway, severity, support vector machine, decision tree, k-Nearest Neighbors, deep learning

Abstract

Railway construction sites are prone to accidents; to be worse, it involves fatality. This is because many factors cannot be controlled due to the hectic working environment. In order to forecast the severity of mishaps at railway construction sites, this study investigates past incidents using machine learning (ML). The study analyzes data from railway construction using k-Nearest Neighbors (k-NN), Decision Trees (DT), Deep Learning (DL), and Support Vector Machines (SVM) implemented in RapidMiner software. ML is used because of its capability to learn about the relationship between each factor and parameter of the incident, thus producing relevant predictions of severity incidents. Finding high-severity occurrences, creating a prediction model, and evaluating the effectiveness of the ML techniques using metrics like accuracy, precision, recall, and F1-score are the objectives. A 70:30 training-testing data split was used, and the results aim to identify the best ML method for predicting incident severity at railway construction sites. SVM and DL are better at predicting the severity of accidents due to their high precision, with both having a 0.91 score for precision. At the same time, DT is favourable for minimising missed critical accidents due to its high recall of 0.89. k-NN shows the most unfavourable performance among these machine learning. This study served as a benchmark for future railway projects, informed mitigation actions and procedures and provided a deeper understanding of potential incidents.

1. Introduction

At the building site, the event and accident were unavoidable. Even with appropriate mitigation measures, the risks can arise at any time and in any location. According to the majority of countries, the construction sector has the greatest prevalence of work-related injuries [1]. The Department of Occupational Safety and Health (DOSH) claims that the construction industry in Malaysia has the highest fatality rate, accounting for almost half of all fatal occupational injuries [2]. Hazard Identification, Risk Assessment, and Risk Control (HIRARC) should be used to investigate and discuss each work sequence in order to appropriately counteract the incident's severity. Health and safety issues are not limited to construction; they can occur at any point in a project's life [3].

It is beneficial to society if it is possible to accurately predict the extent of potential injuries [4]. Creating a thorough predictive assessment model is essential, especially in the Malaysian context, where practitioners continue to use conventional assessment techniques [5]. Thus, this study investigates the prediction of incident severity in railway construction sites using machine learning algorithms. Construction projects are inherently risky due to changing environments and numerous workers. Predicting the severity of potential incidents can significantly benefit safety protocols. In order to predict incident severity and compare the effectiveness of various machine learning algorithms, this study intends to determine the greatest number of incidents and their severity in a railway construction site project. To do this, a model is developed using RapidMiner by implementing several machine learning algorithms, including Support Vector Machine (SVM), Decision Tree (DT), k-Nearest Neighbors (k-NN), and Deep Learning (DL). Large volumes of data can be handled by machine learning-based methods, and more significantly, these methods have the ability to forecast and interpret outcomes. In fact, machine learning techniques have been particularly used to the estimation of accident consequences [4].

1.1 Safety Incidents in Construction

Construction projects are inherently risky due to changing environments and numerous workers. Predicting the severity of potential incidents can significantly benefit safety protocols. With an emphasis on Malaysia, this paper examines recent studies on safety events in building projects as well as machine learning algorithms for predicting incident severity that use SVM, DT, k-NN, and DL

According to DOSH data from 2021, there were 151 fatalities and injuries; 37% occurred in the construction industry, and 28% were manufacturing-related [2]. The safety and well-being of construction site workers are among the most significant health issues [6]. One of the riskiest sectors in the world is construction, and Malaysia has seen a high number of both deadly and non-fatal incidents. Falling objects, falling from heights, and insufficient safety precautions are common causes. The risk is further increased by inadequate safety management and regulatory gaps. The safety and well-being of construction site workers are among the most significant health issues [6]. Inadequate occupational safety and health (OSH) has become a concern in the building industry, which is the most risky with complex procedures that result in many mishaps and fatalities for the public and construction participants [3].

The absence of personal protective equipment was one of the leading contributing causes, and the presence of Safety Site Supervisors (SSS) is mandatory in any situation involving work at heights [7]. A significant percentage of both fatal and non-fatal accidents reported in all industries are caused by Malaysia's construction industry [5]. Compared to supervisors, regular workers are more likely to die in falls [2]. On construction sites, accidents are usually caused by senior management, even though frontline workers' risky activities are the main cause [1]. According to Moshood, work experience, physical health, educational background, professional competence, and emotional intelligence are the five human elements that influence contractor risk attitudes in the Malaysian construction sector [8]. The total number of building incidents in Malaysia from 2014 to 2023 is displayed in Figure 1.

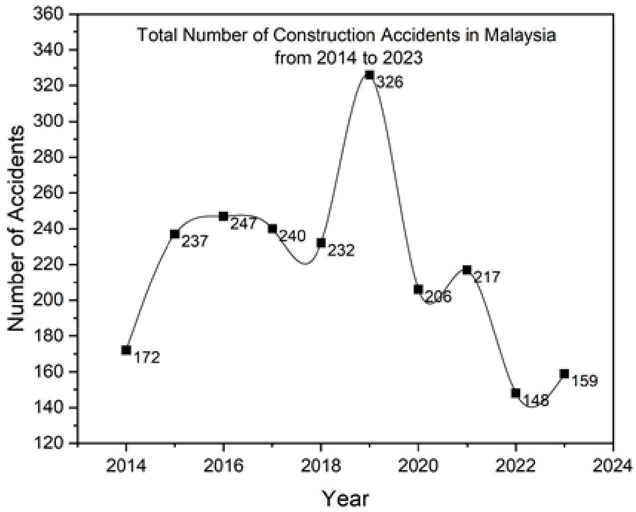

Fig. 1 *Total number of construction accidents in Malaysia from 2014 to 2023 [Statista, 2023]*

Additionally, many safety risks highlighted during the initial research interviews were primarily caused by the main contractors' inadequate supervision [5]. This demonstrates unequivocally that poor communication and human error are major factors in assessing the seriousness of an occurrence [9]. The severity of construction collapse accidents is mostly influenced by three factors: personnel without certifications, project management that disregards safety, and an inability to perform the responsibilities of zonal safety supervision [10]. The most dangerous are tunnels and short-term projects such temporary roads, beam-fabricating yards, track slab yards, mixing stations, yards for processing components and accessories, construction crew stations, and temporary communication stations [11]. Additionally, according to the author, temporary projects ought to be examined and assessed at each phase of the construction process, especially the safety inspection.

1.2 Machine Learning for Incident Severity Prediction

Large amounts of construction data can be analyzed by machine learning algorithms to find trends and forecast the severity of incidents. This makes it possible to allocate resources more effectively and implement preventative safety measures. One of the main objectives of machine learning is to find patterns in data [12]. Because of the massive volumes of data being gathered by contemporary technologies, artificial intelligence algorithms can be utilized to forecast incident outcomes [13]. SVM and DL are commonly used in machine learning [14]. In engineering, stock market forecasting, weather prediction, and medical diagnostics, DL and SVM have proven to be beneficial and perform exceptionally well [14]. To complete all pertinent features, the first phase entails gathering thorough accident data from final incident reports [15]. Cleaning this data is frequently necessary to remove typos, repetitions, and inconsistencies [16-17]

Kaur [18] highlighted five critical reasons for utilising machine learning. First, a strong basis for model training can be established by adequately developing the massive data sets for training. Second, the machine learning algorithms can function efficiently without pre-existing biases because the study does not require the use of logic or commonsense explicit reasoning based on prior information. Thirdly, there is no need for extensive explanations of how the decisions were made, simplifying the implementation process. Furthermore, a phenomena or function becomes stagnant after being taught for a long time, highlighting the significance of continuous data gathering and education to preserve the relevance and accuracy of the model. A proactive approach to safety awareness and process improvement has been fostered by the proposed framework, which has improved communication, safety compliance, and learning within the operation team [19]. Lastly, the task possesses flexible demand constraints, and numerical simulations can offer necessary performance guarantees in case of algorithmic shortcomings, ensuring reliable outcomes even in the presence of potential issues.

Because it determines the best hyperplane to divide data points according to severity, the Support Vector Machine (SVM) is useful for classification problems. The SVM technique aims to identify a hyperplane in an N-dimensional space that classifies the data points [20]. In order to minimize the inaccuracy, the margins are drawn to minimize the distance between the margin and the classes [21]. In the meantime, the dot products in kernel function as an SVM framework are displayed in Figure 2.

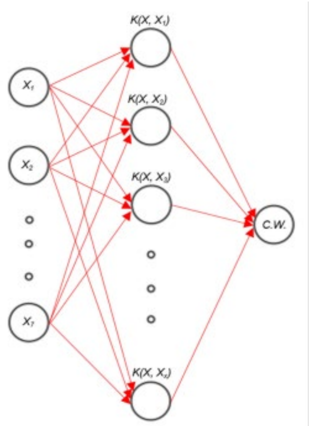

Fig. 2 *The SVM framework (Dot products in kernel function) [14]*

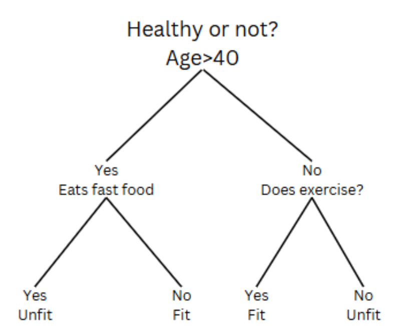

Fig. 3 *Framework of decision tree [21]*

To classify incidents according to characteristics like weather and root cause, Decision Tree (DT) builds a structure like a tree. For classification and regression tasks, a decision tree is a non-parametric supervised learning approach that makes use of a tree-like model. A test on an attribute is represented by each internal node in the tree, and the test's result is represented by each branch. Each node's splitting rule seeks to minimize a selected cost function, usually determined by the error for the selected parameter criterion. Until a stopping requirement is satisfied, like attaining a maximum depth or a minimum degree of cost function improvement, the tree construction process iteratively generates new nodes. Three criteria need to be decided upon to form a tree [22]: the split of each node into son nodes, the accuracy of the classification, and the selection of the final tree for classification. An example of the framework of the decision tree can be seen in Figure 3.

The k-Nearest Neighbors (k-NN) classifies new incidents by comparing them to similar incidents in the past. It does not acquire a model during training, so it fits into the instance-based or lazy learning category. This method employs a memory-based approach. Instead of explicitly learning a model from the training data, it stores all the data points in memory. New instances are compared to these stored examples during prediction to make forecasts.

Deep Learning (DL) uses complex neural networks to learn intricate relationships between data points, potentially leading to highly accurate predictions. DL mimics brain neurons and is associated with machine-based learning algorithms [14]. The two primary characteristics of DL are its capacity to learn and execute complex functions after training, as well as its ability to generalise and come up with a workable solution for unattended data [14]. Figure 4 shows the basic ANN layer with interconnected nodes. The input layer, the first layer in a neural network architecture, is in charge of receiving and allocating the input features. In order to facilitate effective network processing, standardization techniques are frequently used to preprocess the input data and guarantee consistent formatting. The input layer is followed by hidden layers, which may be stacked or single depending on the intricacy of the network. These layers carry out the network's fundamental calculations, using weighted connections and activation functions to derive complex non-linear correlations from the data. Lastly, the output layer contains the neurons that use the data processed from the previous hidden layers to create the network's final predictions.

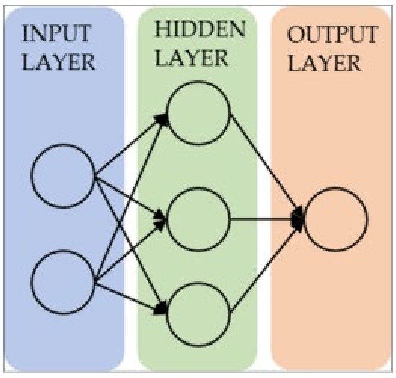

Fig. 4 *The basic ANN layer with interconnected nodes* [14]

Root Mean Square Error (RMSE) is the performance used in the regression (performance) operator because it can be observed to show a more consistent standard deviation of residuals (prediction errors) up to 5%, which approaches actual crack widths, in contrast to Absolute Error (AE) (26–37%) and Prediction Average (PA) (50%). According to [14], who used this method to obtain the lowest error rate for projected crack widths, this study uses a 70:30 training-testing split ratio.

Each machine learning's performance indicators are taken from the body of current literature and utilized for comparison. Accuracy, precision, recall, and F1 score are displayed in equations 1 through 4 [17]. The number of correctly identified observations out of all the observations is known as accuracy. Precision shows the proportion of accurate positive forecasts. The number of positive class observations that the classifier properly predicts is known as recall. The harmonic mean of precision and recall is shown by the F1-score.

$$\frac{Number\ of\ Correct\ Observation}{Number\ of\ Total\ Observation} \text{ x } 100 \tag{1}$$

$$\frac{True\ Positive}{True\ Positive + False\ Positive} \tag{2}$$

$$\frac{True\ Positive}{True\ Positive + False\ Negative} \tag{3}$$

$$2\left(\frac{Recall \text{ x } Precision}{Recall\ +\ Precision}\right) \tag{4}$$

Int. J. of Sustainable Construction Engineering and Tech. Vol. 15 No. 4 (2024) p. 53-63

According to earlier studies, machine learning can greatly enhance building safety procedures. By examining how well these algorithms anticipate the severity of incidents related to railway construction in Malaysia, this study seeks to advance this subject. Therefore, this study utilises incident data from railway constructions in Malaysia. The model considers place, management, employee, weather or surrounding conditions, root causes, likelihood number of deaths and injuries to predict severity. The predicted severity is from four (4) different MLs: SVM, DT, KNN and DL. The predicted severity can be studied to see the performance of each ML and minimise the wrong prediction and the severity of the incident.

2. Methodology

Data on 167 incidents occurred during the railway building project between 2017 and 2023. A format that is appropriate for machine learning algorithms is created from the data. Numerical encoding is used in this work to give categorical attributes distinct numerical values. The effectiveness of many machine learning methods, such as Support Vector Machines (SVM), Decision Trees (DT), k-Nearest Neighbors (k-NN), and Deep Learning (DL), for forecasting the severity of traffic accidents is examined in this study. Following training, the models' accuracy in predicting severity levels is assessed using a different test set [14]. The actual severity of the input data is labeled, and the anticipated severity level is assigned a different attribute. Figure 5 shows the framework model applied for all prediction models from RapidMiner Machine, while Figure 6 shows the testing and training involved in split validation.

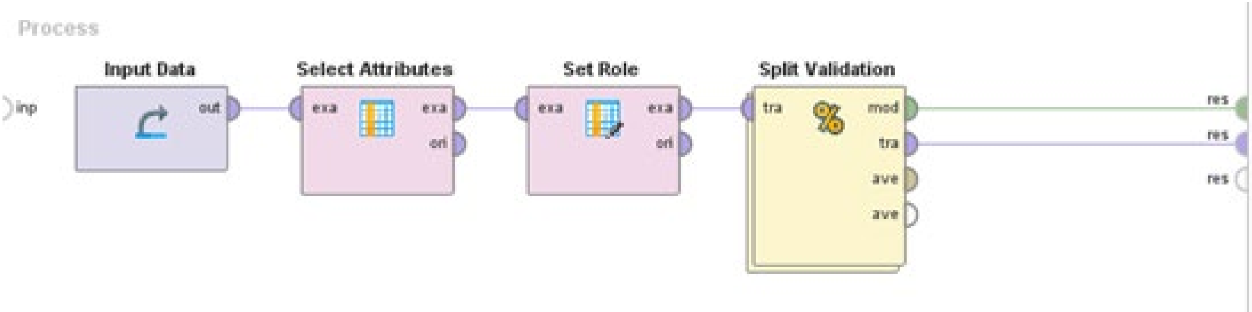

Fig. 5 *The framework model applied for all prediction models from RapidMiner Machine Learning*

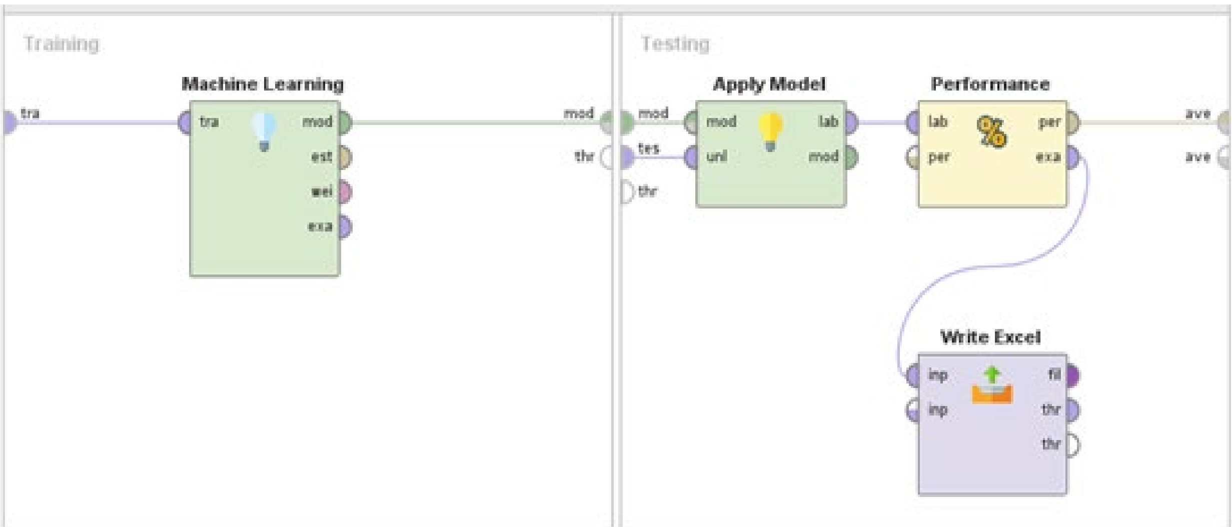

Fig. 6 *Testing and training that involves in split validation*

The parameters being counted into the model are section, sector, incident classification, employee, weather or surrounding condition, root cause, number of injured people, number of people who died, the likelihood of the incident and the severity level. The parameter is decided through progressive discussion with related personnel that can be utilised in the predictive model. Each parameter was assigned a numerical identifier, as RapidMiner could not interpret text but could process numerical values. Table 1 shows the details of each parameter used.

Table 1 *Details of each parameter*

Section	Sector	Incident Classification	Weather or Surrounding Conditions	Root cause	Likelihood of incident	Severity of incident
1	Subgrade	Fatality	Clear skies	Human	1	1
2	Tunnel	Property damage	Cloudy	Environment	2	2
3	Bridge	Lost Time Injury	Rain	Machine	3	3
4	Temporary yard	Occupational Disease	Drizzle rain	Management	4	4
5	Site Area	Environmental	Thunderstorm	Method	5	5
6	Batching Plant	Occupational Poisoning	Fog/Haze	Material		
7	Access to site	Without lost time injury	Windy			
8	Workers Quarters	Dangerous Occurrence	Sufficient lighting			
9	Outside site area	Near miss	Insufficient lighting			
	Base Camp	Medical treatment	Sunny days			
	Main camp	Others	Confined space			
	Site Camp					

Dividing the data into training and testing sets is a standard procedure. There are two uses for the training set: (1) building the model and (2) providing the foundation for parameter estimation, model comparison, and other development tasks. The testing set provides an objective evaluation of the model's performance and is kept apart until the final assessment. The 70:30 ratio is used in this study for both testing and training.

Using the test set during model development can lead to biased results, potentially inflating the model's output due to overfitting. To prevent this, the data is split to maximise training data accuracy while avoiding overfitting and underfitting. RapidMiner's split ratio is crucial in evaluating model efficacy and accuracy. A successful model works well on unseen data because it generalizes well. In addition to testing the model on a sample fraction to evaluate performance of unseen instances, an ideal split ratio guarantees that the model is trained on enough data to identify patterns and generate precise predictions.

After the result had been extracted, the data was compared between actual and predicted data to classify the correct observations, false positives, and false negatives and the performance metrics were calculated as stated in Equation 1 to Equation 4. Data must be categorized into three groups: false positives (where the predicted data is wrong but underestimates the severity compared to the actual data), false negatives (where the predicted data is wrong but overestimates the severity compared to the actual data), and correct observations (where the actual data matches the predicted data). Calculating accuracy, precision, recall, and the F1-Score requires this classification.

3. Result and Discussion

From the 167-incident data, the ratio of 70:30 was used to split it according to training and testing. The testing data comprised of 50 datasets, while the training dataset consisted 112, which was used by machine learning to learn the algorithm of the incident from each parameter. From these 50 datasets, the predicted severity was made, and the correct observation, false positive and false negative, were considered for the performance metrics.

3.1 Machine Learning for Incident Severity Prediction

The performance metrics are accuracy, precision, recall and F1-score. SVM shows correct observations of 40 out of 50, with four false positives and six false negatives. DT shows correct observations of 40 out of 50, with seven false positives and three false negatives. k-NN shows correct observations of 34 out of 50, with six false positives and ten false negatives. Meanwhile, DL shows correct observations of 41 out of 50, with four false positives and five false negatives. Table 2 shows the severity of the project.

Table 2 *Description for severity (HIRARC for the project 2021)*

Severity	Description
1	First Aid Required
2	Minor injury/<4 days lost time
3	Major injury/> four days lost time
4	Permanent disability
5	Fatality/multiple fatality

Figure 7 shows the number of incidents according to severity level for railway construction projects from 2017 until 2023. Health issues primarily cause incidents with a severity level of five, delayed emergency responses, vehicle accidents due to loss of control, inadequate traffic management, and communication breakdowns during work. These incidents are strongly linked to management and human factors. Workers experiencing health problems, such as heart attacks, may lose control and get injured, and delayed medical attention can worsen outcomes. Vehicle accidents often result from driver errors and poor traffic management. Misunderstandings or a lack of communication between workers and supervisors also contribute to unsafe situations. To improve safety, workplaces should promote worker health, establish clear emergency response protocols, implement effective traffic management practices, and ensure clear communication throughout the workday.

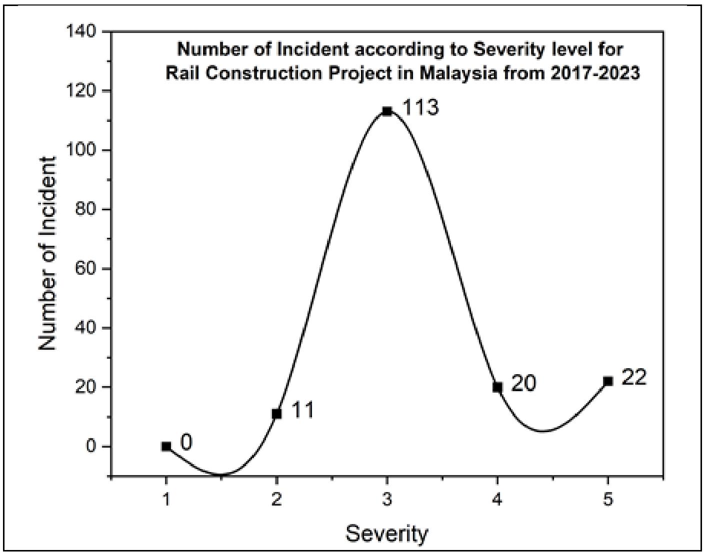

Fig. 7 *Number of incidents according to severity level for rail construction project in Malaysia from 2017 to 2023*

3.2 Performance Metrics for Each Machine Learning

Evaluation of the machine learning model in Figure 8 using metrics like accuracy, precision, recall, and F1-score revealed exciting insights. Despite achieving the highest precision of 0.91 and accurately identifying actual high-severity incidents, SVM exhibited a lower accuracy of 0.8. This suggests prioritising the positive class (likely high severity) with high confidence, leading to more true positives and some false positives. Precision becomes particularly important when false positives are costly, as in medical diagnosis. In imbalanced datasets, where most incidents might be low severity, accuracy can be misleading. A model biased towards the majority class can achieve high accuracy but fails to classify the minority class (high-severity) incidents correctly. Therefore, depending on the cost of false positives and negatives in your specific severity prediction application, the F1-score, which balances precision and recall, might be a more suitable metric for a more comprehensive evaluation.

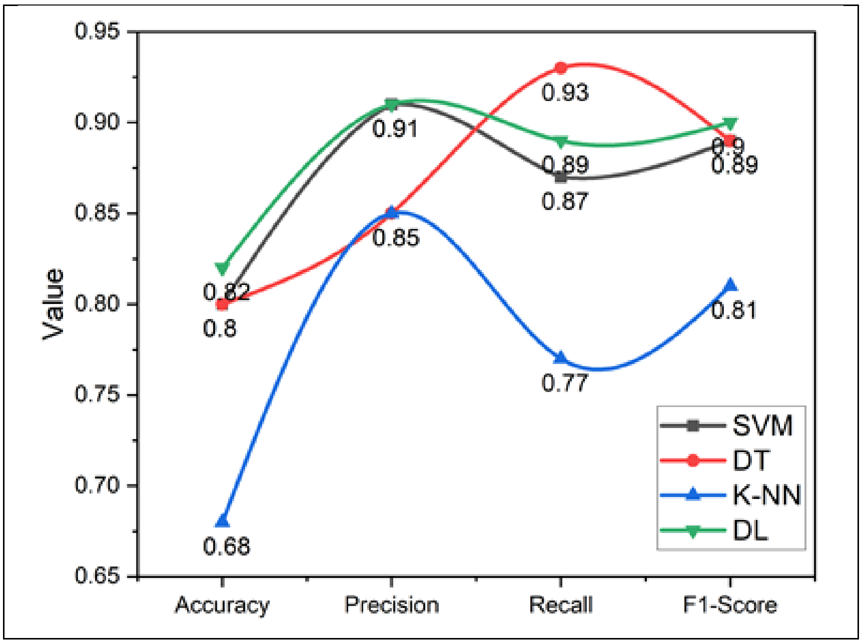

Fig. 8 *Machine learning performance result*

The DT model was evaluated using evaluation criteria such as accuracy, precision, recall, and F1-score (formulas previously provided). Notably, DT's greatest recall of 0.93 demonstrated how well it was able to extract positive instances—likely high-severity incidents—from the training set. The recursive partitioning method used by DT, which produces feature space regions enriched with positive occurrences, is the source of its strength. The model's recall quantifies its capacity to accurately classify every real positive case.

DT, on the other hand, had a lower accuracy score (0.8). This might be explained by possible overfitting, especially in situations with complicated or noisy data. When a model overfits, it captures noise or outliers in the training data, which reduces accuracy and makes it more difficult for the model to generalize to new data. Overall accuracy may be impacted by DT's complex decision boundaries, which may also lead to erroneous predictions in some areas.

Using the previously given formulas for accuracy, precision, recall, and F1-score, the performance of the k-NN model was assessed. Although it had a lesser accuracy of 0.68, k-NN had the highest precision of 0.85. This is explained by the data's class imbalance, in which there are substantially more incidents in one class (presumably low-severity episodes) than in the other. In such cases, k-NN might prioritise correctly classifying instances from the majority class (high precision) at the expense of identifying positive instances (high-severity incidents) from the minority class (low recall). The dominance of the majority class during k-NN's decision-making process can potentially lead to the misclassification of minority class instances.

Additional factors influencing performance include the chosen distance metric and K value. Inappropriate choices can lead to misinterpretations of the underlying data distribution and contribute to the imbalanced precision and recall observed here. Furthermore, k-NN relies on local information from nearby instances to determine decision boundaries. This local approach can result in boundaries that closely follow the distribution of the majority class, especially in areas with interspersed positive and negative instances. In these regions, the local decision-making process might miss positive instances that lack sufficient positive neighbours, leading to a lower recall.

The DL model was evaluated using evaluation measures such as accuracy, precision, recall, and F1-score (formulas previously provided). DL achieved the highest precision with 0.91 but exhibited a lower accuracy of 0.82. This prioritisation of precision can be attributed to several factors. Deep learning models excel at discriminating between positive and negative instances, making them well-suited for scenarios where false positives are particularly costly, such as medical diagnosis. Imbalanced class distributions, where one class (likely low-severity incidents) dominates the data, can also influence DL to prioritise accurately predicting the majority class, potentially sacrificing overall accuracy. Additionally, deep learning models might overfit, particularly when dealing with sparse or noisy data. While overfitting increases precision on training data, it reduces accuracy by impeding generalization to new data. Nonetheless, deep learning architectures' adaptability enables them to identify intricate connections in the data, producing accurate forecasts for certain data points or geographical areas.

3.3 Comparison of Performance Metrics in Machine Learning

Because Deep Learning (DL) can discover intricate patterns and relationships in vast, complicated datasets, it achieved the best accuracy (0.82). Because of this, DL is ideally suited to capturing the subtle characteristics linked to the severity of accidents at railroad construction sites. The comparison of each machine learning performance metric is displayed in Figure 9.

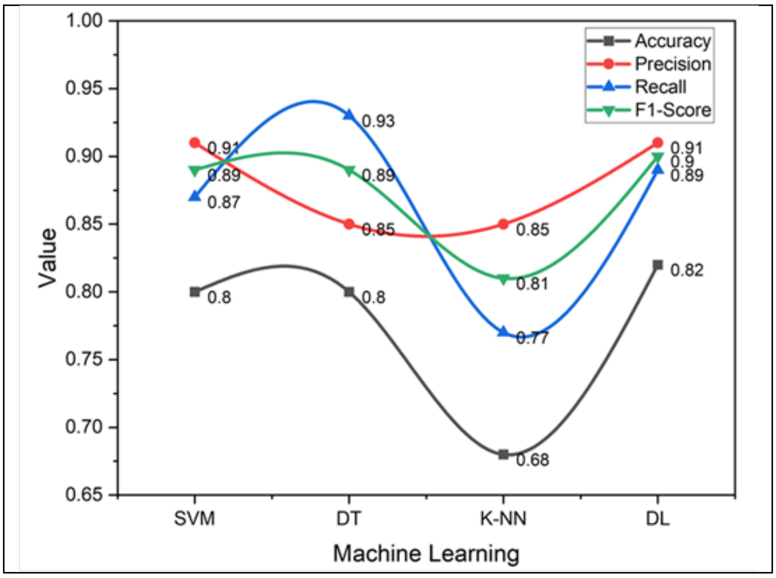

Fig. 9 *Comparison of each performance metric for each machine learning*

The maximum precision was attained with DL and Support Vector Machines (SVM) (0.91). They are able to capture the intricacies of accident severity data by efficiently separating several classes in high-dimensional data thanks to their capacity to learn complex decision boundaries. When working with unbalanced datasets when one class (such as serious accidents) is substantially smaller, precision is essential. In order to address these imbalances, DL and SVMs can modify their decision bounds to give precise minority class instance identification priority. In comparison to k-Nearest Neighbors (k-NN) and Decision Trees (DT), they are also more resilient to noise and outliers.

DT achieved the highest recall (0.93) due to its ability to create complex decision boundaries that adapt to underlying data patterns. This flexibility allows DT to capture nuanced relationships between features and accident severity, leading to higher recall of positive instances (likely high-severity accidents). k-NN exhibited the lowest recall of 0.77 as it relies on local similarity measures and does not explicitly model decision boundaries. This limitation hinders its ability to capture complex relationships, especially in imbalanced datasets where the minority class (severe accidents) might be underrepresented.

With the greatest F1-score (0.9), SVM demonstrated a good trade-off between recall and precision. This strength results from SVM's ability to capture complex decision boundaries because to its efficacy in addressing non-linear interactions through kernel functions. Due to possible issues with specific non-linearities or feature interactions, especially when dealing with noisy data or restricted tree depth, DT got the lowest F1-score (0.81). For unbalanced datasets, the F1-score is crucial, and SVM's strong F1 score is a result of its capacity to modify margins and decision boundaries for the minority class. The identical F1-score (0.89) was attained by k-NN and DL, indicating their trade-off between recall and precision.

4. Conclusion

The effectiveness of many machine learning models for estimating the seriousness of accidents at railroad construction sites was compared in this study. Health problems and car accidents are frequent high-severity incidences in this project. Metrics like accuracy, precision, recall, and F1-score were used to assess the models. Because Deep Learning (DL) can recognize intricate patterns in big datasets, it was able to attain the best accuracy of 0.82. Models like Support Vector Machines (SVM) and DL, which had the maximum precision of 0.91 and 0.91, respectively, are better for unbalanced class distributions because they may give priority to accurately identifying the minority class (such as serious accidents). Because of their high recall score (0.93), Decision Trees (DT) may be a useful option when working with unbalanced data and trying to achieve a high recall of positive cases (serious accidents). With the greatest F1 score of 0.9, SVM, on the other hand, emerged as the overall leader, demonstrating a performance that balanced recall and precision. This shows how well SVM handles imbalanced datasets and non-linear relationships. The priorities and features of the railway accident severity prediction challenge determine which model is best.

Acknowledgement

The authors express their deep gratitude to Universiti Tun Hussein Onn Malaysia for their unwavering support in facilitating this research work. The authors also thank the colleagues from UTHM who provided insight and expertise that greatly assisted the research.

Conflict of Interest

Authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

All authors equally contributed to this manuscript.

References

[1] C. Wu, D. Fang, and N. Li(2015), Roles of owners' leadership in construction safety: The case of high-speed railway construction projects in China, *Int. J. Proj. Manag.*, vol. 33, no. 8, pp. 1665–1679, doi: 10.1016/j.ijproman.2015.07.005. https://doi.org/10.1016/j.ijproman.2015.07.005

[2] A. Zermane, M. Z. M. Tohir, M. R. Baharudin, and H. M. Yusoff(2023), Investigating patterns of workplace fatal fall injuries: Case study of Malaysia, *J. Safety Res.*, vol. 85, pp. 492–506, doi: 10.1016/j.jsr.2023.05.003. https://doi.org/10.1016/j.jsr.2023.05.003

[3] N. M. Saifullah and F. Ismail(2012), Integration of Occupational Safety and Health during Pre-construction Stage in Malaysia, *Procedia - Soc. Behav. Sci.*, vol. 35, no. December 2011, pp. 603–610, doi: 10.1016/j.sbspro.2012.02.127. https://doi.org/10.1016/j.sbspro.2012.02.127

[4] R. Zhu, X. Hu, J. Hou, and X. Li(2021), Application of machine learning techniques for predicting the consequences of construction accidents in China, *Process Saf. Environ. Prot.*, vol. 145, pp. 293–302, doi: 10.1016/j.psep.2020.08.006. https://doi.org/10.1016/j.psep.2020.08.006

[5] H. Sadeghi, S. R. Mohandes, M. R. Hosseini, S. Banihashemi, A. Mahdiyar, and A. Abdullah(2020), Developing an ensemble predictive safety risk assessment model: Case of Malaysian construction projects, *Int. J. Environ. Res. Public Health*, vol. 17, no. 22, pp. 1–25, doi: 10.3390/ijerph17228395. https://doi.org/10.3390/ijerph17228395

[6] O. Ugur, A. A. Arisoy, M. Can Ganiz, and B. Bolac(2021), Descriptive and prescriptive analysis of construction site incidents using decision tree classification and association rule mining, *2021 Int. Conf. Innov. Intell. Syst. Appl. INISTA 2021 - Proc.*, pp. 1–6, doi: 10.1109/INISTA52262.2021.9548427. https://doi.org/10.1109/INISTA52262.2021.9548427

[7] A. Zermane, M. Z. Mohd Tohir, M. R. Baharudin, and H. Mohamed Yusoff(2022), Risk assessment of fatal accidents due to work at heights activities using fault tree analysis: Case study in Malaysia, *Saf. Sci.*, vol. 151, no. March, p. 105724, doi: 10.1016/j.ssci.2022.105724. https://doi.org/10.1016/j.ssci.2022.105724

[8] T. D. Moshood, A. Q. Adeleke, G. Nawanir, and F. Mahmud(2020), Ranking of human factors affecting contractors' risk attitudes in the Malaysian construction industry, *Soc. Sci. Humanit. Open*, vol. 2, no. 1, p. 100064, doi: 10.1016/j.ssaho.2020.100064. https://doi.org/10.1016/j.ssaho.2020.100064

[9] H. Lan, X. Ma, W. Qiao, and W. Deng(2023), Determining the critical risk factors for predicting the severity of ship collision accidents using a data-driven approach, *Reliab. Eng. Syst. Saf.*, vol. 230, no. October 2022, p. 108934, doi: 10.1016/j.ress.2022.108934. https://doi.org/10.1016/j.ress.2022.108934

[10] X. Luo, X. Li, Y. M. Goh, X. Song, and Q. Liu(2023), Application of machine learning technology for occupational accident severity prediction in the case of construction collapse accidents, *Saf. Sci.*, vol. 163, no. August 2022, p. 106138, doi: 10.1016/j.ssci.2023.106138. https://doi.org/10.1016/j.ssci.2023.106138

[11] C. Lu and C. Cai(2019), Challenges and Countermeasures for Construction Safety during the Sichuan–Tibet Railway Project, *Engineering*, vol. 5, no. 5, pp. 833–838, doi: 10.1016/j.eng.2019.06.007. https://doi.org/10.1016/j.eng.2019.06.007

[12] M. S. Kumar, P. Neeraja, R. Amulya, K. Swarna, and T. S. P. Reddy(2023), Severity Level Prediction and Finding Cause in Flight Accidents Using Machine Learning, *2023 Int. Conf. Sustain. Emerg. Innov. Eng. Technol. ICSEIET 2023*, pp. 565–568, doi: 10.1109/ICSEIET58677.2023.10303615. https://doi.org/10.1109/ICSEIET58677.2023.10303615

[13] F. Alsakka, Y. Mohamed, and M. Al-Hussein(2021), Evaluation and Comparison of the Performance of Artificial Intelligence Algorithms in Predicting Construction Safety Incidents, *Proc. Int. Symp. Autom. Robot. Constr.*, vol. 2021-November, no. Isarc, pp. 568–575, doi: 10.22260/isarc2021/0077. https://doi.org/10.22260/isarc2021/0077

[14] R. Ganasan, C. G. Tan, Z. Ibrahim, F. M. Nazri, M. M. Sherif, and A. El-Shafie(2021), Development of crack width prediction models for rc beam-column joint subjected to lateral cyclic loading using machine learning, *Appl. Sci.*, vol. 11, no. 16, doi: 10.3390/app11167700. https://doi.org/10.3390/app11167700

[15] M. F. Labib, A. S. Rifat, M. M. Hossain, A. K. Das, and F. Nawrine(2019), Road Accident Analysis and Prediction of Accident Severity by Using Machine Learning in Bangladesh, *2019 7th Int. Conf. Smart Comput. Commun. ICSCC 2019*, pp. 1–5, doi: 10.1109/ICSCC.2019.8843640. https://doi.org/10.1109/ICSCC.2019.8843640

[16] S. I. B. Ginting and M. Iqbal(2022), Application of Data Mining to Predict Work Accident Rates using Rapid Miner, *J. Mantik*, vol. 5, no. 4, pp. 2694–2701.

[17] S. Ahmed, M. A. Hossain, M. M. I. Bhuiyan, and S. K. Ray(2021), A Comparative Study of Machine Learning Algorithms to Predict Road Accident Severity, *Proc. - 2021 20th Int. Conf. Ubiquitous Comput. Commun. 2021 20th Int. Conf. Comput. Inf. Technol. 2021 4th Int. Conf. Data Sci. Comput. Intell. 2021 11th Int. Conf. Smart Comput. Networking, Serv. IUCC/CIT/DSCI/SmartCNS 2021*, pp. 390–397, doi: 10.1109/IUCC-CIT-DSCI-SmartCNS55181.2021.00069. https://doi.org/10.1109/IUCC-CIT-DSCI-SmartCNS55181.2021.00069

[18] M. Kaur, A. K. Shukla, and S. Kaur(2021), An Introduction to Machine Learning in a Nutshell, *Proc. 2021 10th Int. Conf. Syst. Model. Adv. Res. Trends, SMART 2021*, pp. 17–22, doi: 10.1109/SMART52563.2021.9676315. https://doi.org/10.1109/SMART52563.2021.9676315

[19] J. B. Martins, G. Carim, T. A. Saurin, and M. F. Costella(2022), Integrating Safety-I and Safety-II: Learning from failure and success in construction sites, *Saf. Sci.*, vol. 148, no. September 2021, doi: 10.1016/j.ssci.2022.105672. https://doi.org/10.1016/j.ssci.2022.105672

[20] Y. Ledmaoui, A. El Fahli, A. Elmaghraoui, M. El Aroussi, R. Saadane, and A. Chehri(2023), Machine Learning Techniques Using the Rapid Miner Tool for Solar Production Forecasting, *Proc. - 2023 Int. Conf. Digit. Age Technol. Adv. Sustain. Dev. ICDATA 2023*, pp. 120–125, doi: 10.1109/ICDATA58816.2023.00030. https://doi.org/10.1109/ICDATA58816.2023.00030

[21] M. Kumar, S. Ali Khan, A. Bhatia, V. Sharma, and P. Jain(2023), A Conceptual introduction of Machine Learning Algorithms, *2023 1st Int. Conf. Intell. Comput. Res. Trends, ICRT 2023*, pp. 1–7, doi: 10.1109/ICRT57042.2023.10146676. https://doi.org/10.1109/ICRT57042.2023.10146676

[22] A. Samimi, H. Razi-Ardakani, and A. Nohekhan(2017), A Comparison between Different Data Mining Algorithms in Freight Mode Choice, *Am. J. Appl. Sci.*, vol. 14, no. 2, pp. 204–216, doi: 10.3844/ajassp.2017.204.216. https://doi.org/10.3844/ajassp.2017.204.216