Journal of Design and Built Environment
E-ISSN: 2232-1500 |Print ISSN: 1823-4208
https://ejournal.um.edu.my/index.php/jdbe
Special Issue V, 2025, 69-86

Moment-Rotation Characteristics Prediction Models for Unique Boltless Steel Connections Using Machine Learning

Reventheran Ganasan[1,2*], Chee Ghuan Tan[3], Muhammad Naiman Arimi Arifin[3], Nor Hafizah Ramli Sulong[4], Mustapha Kamil Omran[1,2], Ahmed El-Shafie[3], Anies Faziehan Zakaria[5]

[1]Department of Transportation Engineering Technology, Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia, 84600, Pagoh, Johor, Malaysia.
[2]Industry Centre of Excellence for Railway (ICOE-REL) - Satellite, Universiti Tun Hussein Onn Malaysia, 86400 Pagoh, Johor, Malaysia.
[3]Department of Civil Engineering, Faculty of Engineering, Universiti Malaya, 50603 Kuala Lumpur, Malaysia.
[4]School of Civil & Environmental Engineering, Faculty of Engineering, Queensland University of Technology, George St, Brisbane, Queensland 4000, Australia.
[5]Department of Engineering Education, Fakulti Kejuruteraan dan Alam Bina, Universiti Kebangsaan Malaysia, Lingkungan Ilmu, 43600 Bangi, Selangor, Malaysia.

*Corresponding Author: reven@uthm.edu.my, tancg@um.edu.my

Article info:

Submission date: 11th September 2024
Acceptance date: 12th December 2024

Keywords:
M-θ, steel pallet racket, boltless steel connections, rapid miner machine learning

ABSTRACT

Beam-to-column connections (BCCs) in pallet rack structures are used for storing goods in industrial buildings, warehouses, and super-stores. BCCs must be easily demountable and reassembled to accommodate changing requirements over time. Common experimental tests for evaluating connection behaviour are expensive and time-consuming, so this study developed three prediction models using different algorithms to assess the moment-rotation behaviour of different connection types. The models were based on Support Vector Machine (SVM), Deep Learning (DL), and Decision Tree (DT) algorithms and trained using 70:30 split ratios, with further testing of 60:40 and 80:20 ratios. The models were evaluated using root mean square error, mean absolute error, and relative coefficient. The modified 60:40 DT Least Square model outperformed the other models in predicting moment-rotation behaviour, with consistent performance across all split ratios. The SVM Radial model performed poorly due to classification errors, and the DL Rectifier model made inconclusive predictions due to small sample size. The study highlights the accuracy and feasibility of various algorithm techniques in predicting BCC behaviour, enabling cost-effective and efficient testing of connections in pallet rack structures.

1.0 INTRODUCTION

Modern production and distribution requirements in competitive markets are met primarily using racking systems. Engineers face a daunting task when selecting storage equipment due to the many alternatives available. Conventional pallet racking options include selective/adjustable racks, double-deep, drive-through layouts, live pallet storage, push-back, and mobile storage systems [1]. These racks, self-supporting thin-walled steel structures, can bear substantial vertical and lateral stresses. To allow users to adjust configurations as requirements change, connections in racking systems must be removable, disqualifying bolted and welded connections. Cold-formed, boltless, semi-rigid connections with beam-to-column connections (BCC) have become prevalent in the pallet manufacturing industry due to their low production and assembly costs and lower material expenditures [2]. Pallet rack systems, used for low-density storage or when all commodities must be accessible, utilise less floor space than other systems. The upright column frames support vertical loads, while horizontal beams are supported by column frames perpendicular to the column planes. A pallet, stacked and transported by a forklift, facilitates item transport. The beam-end connector (BEC), made from hot-rolled steel and equipped with hooks or tab connectors, provides a boltless connection method based on the upright's perforation pitch. Tabs, produced and punched out of the beam-end connector, are an integrated component of the BEC. This results in a semi-rigid but stable structure due to upright holes on the vertical walls [3-5]. The BEC is crucial for down-aisle stability, supporting beams and ensuring the rack frame's sway stability, as racks are typically not braced in the down-aisle orientation. Consequently, pallet rack systems, with semi-rigid connections between beams and columns, are susceptible to structural collapse due to lateral stresses.

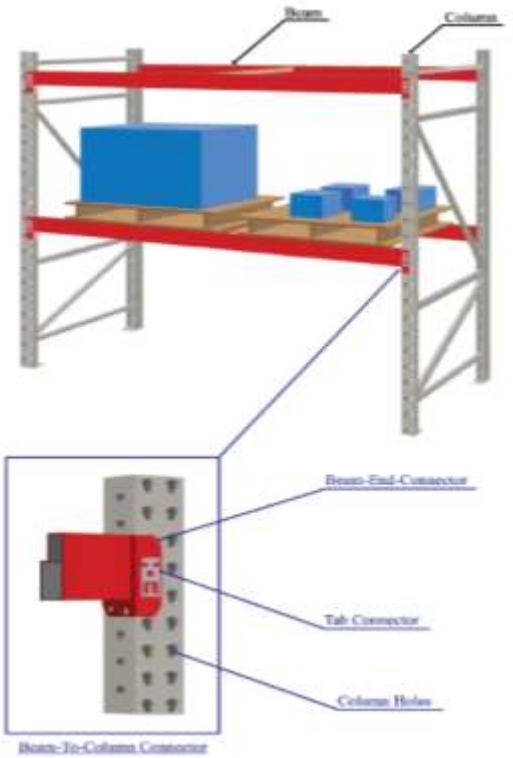


Figure 1. Typical steel pallet racking system.

Research on boltless steel connections is limited, particularly under seismic and fire conditions [6-7], and numerical analysis [8]. Experimental testing is commonly used to predict connection performance factors and develop universal design strategies. However, experimentation is costly, time-consuming, and hard to replicate. Previous studies show that finite element modelling (FEM) often fails to accurately capture failure modes seen in experiments [9]. While FEM predictions align well with experimental results in the linear elastic range, they overestimate the ultimate moment in the plastic stage [10]. Bernuzzi and Castiglioni noted that predicting the stiffness and strength of boltless connections analytically is not feasible due to their diversity, leading to the recommendation of testing in international rack design standards [11]. Aguirre's studies on boltless connections under static and cyclic stresses found that the connecting members determine the failure

load [12]. The variability in the geometrical aspects of boltless connections complicates the creation of a general analytical formula for predicting moment-rotation (M-θ) characteristics. Recent consensus suggests using artificial intelligence (AI) and soft computing (SC) technologies to address design issues, reducing the need for costly experimental testing [13-14]. This study aimed to develop AI-based prediction models for the M-θ characteristics of boltless steel connections. The framework was created using RapidMiner, employing three machine learning (ML) algorithms: Support Vector Machine (SVM), Deep Learning (DL), and Decision Tree (DT). The model was based on experimental data with a 70:30 training-to-testing split. Models were evaluated using the correlation coefficient (R2), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). DT Least Square showed the best prediction capacity, prompting further analysis of how different data splitting ratios (60:40, 70:30, and 80:20) affect performance. This study considered three input parameters: column thickness, beam depth, and boltless connection depth, identifying eight distinctive boltless steel connections. The output focused on the applicability of the three ML algorithms in developing regression models to predict the M-θ characteristics of boltless connections.

2.0 MATERIALS AND METHODS

This study utilised the outcomes of 16 tests performed by Shah et al. [6] to assess the advancement in SPR and BCC performance under static loading. Two trials of each set of combinations were done by altering the column's thickness, the beam's depth, and the BEC's depth. Table 1 shows the material properties of the test specimen. A constant column height of 500 mm was maintained throughout the experiment. The double cantilever test method was used for investigations. At a rate of 3 mm/min and with a hydraulic actuator capable of 50 kN of force, the machine exerted pressure on the top of the column until it failed. Table 2 provides the sample size of 8020 for each combination. These data were used as primary inputs to generate the outcome of predicted rotation.

Table 1. Material properties of test specimens.

Member	Young Modulus (E) (GPa)	Yield Strength (f_y) (MPa)	Ultimate Strength (f_u) (MPa)
Column	210	459	575
Beam		353	497
Beam-end-connector		263	365

Table 2. Number of sample data for each combination

Sample	Column Thickness (mm)	Depth of Beam (mm)	Depth of BEC (mm)	Number of Sample Data
A	2	92	205	1534
B	2	110	205	1074
C	2	125	255	295
D	2	150	255	768
E	2.6	92	205	1823
F	2.6	110	205	745
G	2.6	125	255	853
H	2.6	150	255	928

2.1. The Neural Network Configuration

In this research endeavor, RapidMiner, a versatile data science platform, harnessed the prowess of three distinct yet potent machine learning models: Support Vector Machine (SVM), Deep Learning (DL), and Decision Tree (DT). Each of these models represents a unique approach to predictive analytics, offering varied capabilities suited for tackling diverse challenges within the dataset under investigation. By leveraging the combined strengths of Support Vector Machine, Deep Learning, and Decision Tree models within RapidMiner, this research aimed to harness a comprehensive arsenal of analytical tools to extract valuable insights, discern intricate patterns, and make accurate predictions, thereby advancing the understanding of the phenomena under study and facilitating informed decision-making.

2.1.1. The Neural Network Configuration

SVM is a collection of closely related supervised learning techniques used for classification and regression [15]. According to Boswell [16], SVM produces hyper-planes or sets of hyper-planes in high-dimensional or infinite-dimensional space. The aim is to develop a function approximating the mapping from an input domain to real numbers based on a training sample regarding both the positive and negative sides. Models with the best fit have the closest data point to their hyperplane and are within the margin of tolerance. The challenge in SVM is to find a function f that returns the best fit:

$$f(w_1, \dots, w_n, b) = y = w \cdot x + b + \varepsilon \quad (1)$$

where w represents coefficients and b is the intercept

The kernel function in SVMs allows for the transformation of the input space into a linear representation in the higher-dimensional space known as the feature space [17]. This is accomplished by adding extra dimensions and then finding an optimum boundary between the various outputs based on these transformations. Figure 2 illustrates the architecture of an SVM algorithm for regression. To make a prediction, the input pattern is put into a feature space that can be mapped. Kernel functions are then used to calculate the final results using the training patterns (support vectors). Finally, the weights of the products are used to add them all up. This, together with the constant term, yields the final regression result.

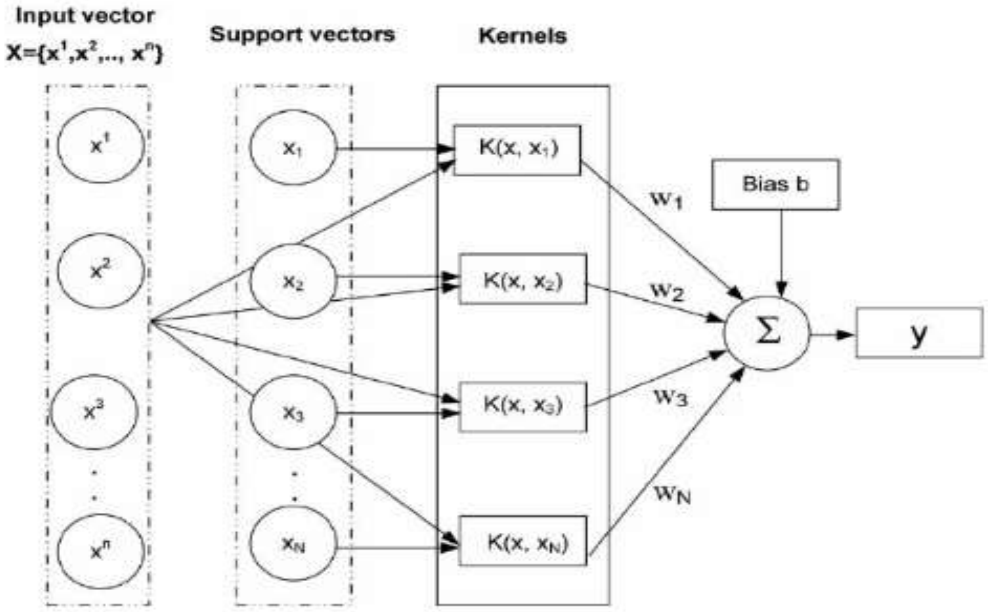


Figure 2. Support vector machine's algorithm

2.1.2. Deep Learning (DL)

Derived initially from the artificial neural network (ANN), DL is a technique that may be used to intelligently analyse data [18]. Deep learning assures the replacement of handmade features with effective approaches for unsupervised or semi-supervised learning and hierarchical feature extraction [19]. Mathew et al. [20] indicated that when dealing with unstructured data, deep learning gets increased power and flexibility owing to its capacity to analyse a vast number of characteristics. The more data there is, the better trained they are. DL's "deep" part involves layering neurons in a network of arbitrary sizes and stacking them on top of one another [21]. Figure 3 depicts the deep learning framework's basic structure. The data is transmitted through a series of layers, each of which can successfully extract characteristics and transmit them to the subsequent layer. Starting with the most basic features, subsequent layers build upon one another to form a complete depiction.

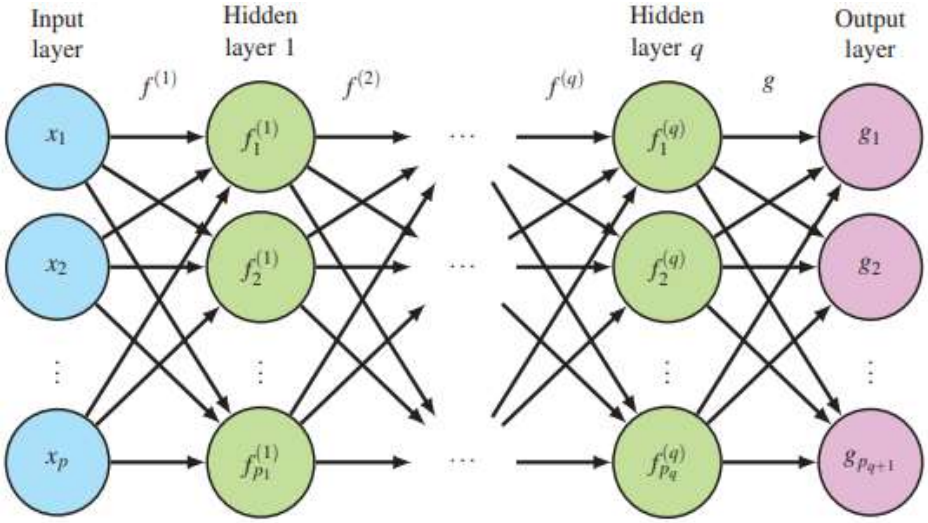


Figure 3. Deep learning's framework

Placing one function to work with the output of other functions is called "function composition". The function composition is written as: *g o f(x) = g(f(x))*, where *g o f* is interpreted as "*g* constructed of *f*" or "*g* of *f*". The function *f(x)* receives an input value *x* and returns an output value *f* when combining g with *f(x)*. Next, *f(x)* is fed to the function *g*, which gives the final value *g f(x)*. Most DL models are composed of a large number of multivariate vector-valued functions, each of which contains parameters that may be optimized to improve the model's ability to predict precisely [22].

2.1.3. Decision tree (DT)

According to Rokach et al. [23], DT is a classifier expressed as a recursive partition of the instance space. Geometrically, DT with numerical characteristics may be seen as a collection of orthogonal hyperplanes, each pointing in the opposite direction as the axes. Once the machine is trained, it begins to predict and decide when new data is given to it. The nodes consisted of the first node, namely the Root Node, which characterizes the entire sample and may be split into other nodes. Following that are Interior Nodes, which reflect the data's attributes and their branches' decision-making rules. Lastly, the final result is represented by the Leaf Nodes. By answering True/False questions, a specific data point is traversed through the entire tree until it reaches the leaf node. A leaf node's average dependent variable value is used to make the final prediction. For the regression model, Leaf nodes may contain either explicit concept values or a function that computes the continuous value of the target attribute [24-25]. When a dataset is submitted to a decision tree inducer, the algorithm builds a decision tree automatically. Typically, the objective is to minimize the generalization error to discover the ideal decision tree. Figure 4 depicts the decision tree's components.

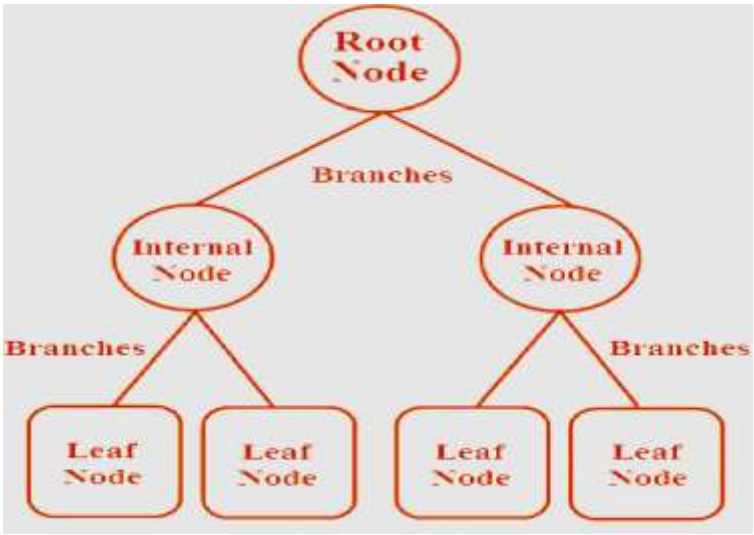


Figure 4. Decision tree's component

2.2. Framework of prediction model

Each model within the study adheres to a unified predictive model framework, differentiated primarily by the predictive operator employed. Initially, experimental data encompassing parameters such as column thickness, beam depth, and depth of the bolted end-plate connection (BEC) were ingested into RapidMiner. Subsequently, within the Data Editor, the attributes corresponding to "Moment (kNM)" and "Rotation (Radians)" were customized, with designated roles as "Label" and "Prediction" respectively. A Split Validation operator was then engaged, employing an automated relative split ratio of 0.7 to segregate the dataset for training and validation purposes. For model training, the attribute denoted as "Rotation (Radians)" was selected with the target role of "Label" and subsequently normalized via a range transformation spanning 0 to 1. The processed data was then fed into a specific predictive operator-based model tailored to each model variant. Utilizing respective machine learning algorithms, prediction models were constructed based on the training data subset. Subsequently, these trained models were applied to the test dataset utilizing the Apply Model operator to predict the M-θ behavior of boltless steel connections.

Evaluation of regression performance, based on the experimental M-θ behavior, was conducted utilizing "Performance" operators, employing three key evaluation metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). The predicted values of M-θ behavior were then exported onto a designated Excel file for further analysis. Leveraging the projected rotation dataset, a comprehensive examination was conducted to discern the correlation between predicted and experimental rotation values. An illustrative framework model for the Decision Tree (DT) variant is depicted in Figure 5, exemplifying the operational process. Subsequently, based on prediction accuracy, the model exhibiting the highest predictive capability was identified for further investigation. The impact of splitting ratios on training-testing distributions (60:40 and 80:20) was then scrutinized through the utilization of the "Split Validation" operator, facilitating a more granular exploration of model performance.

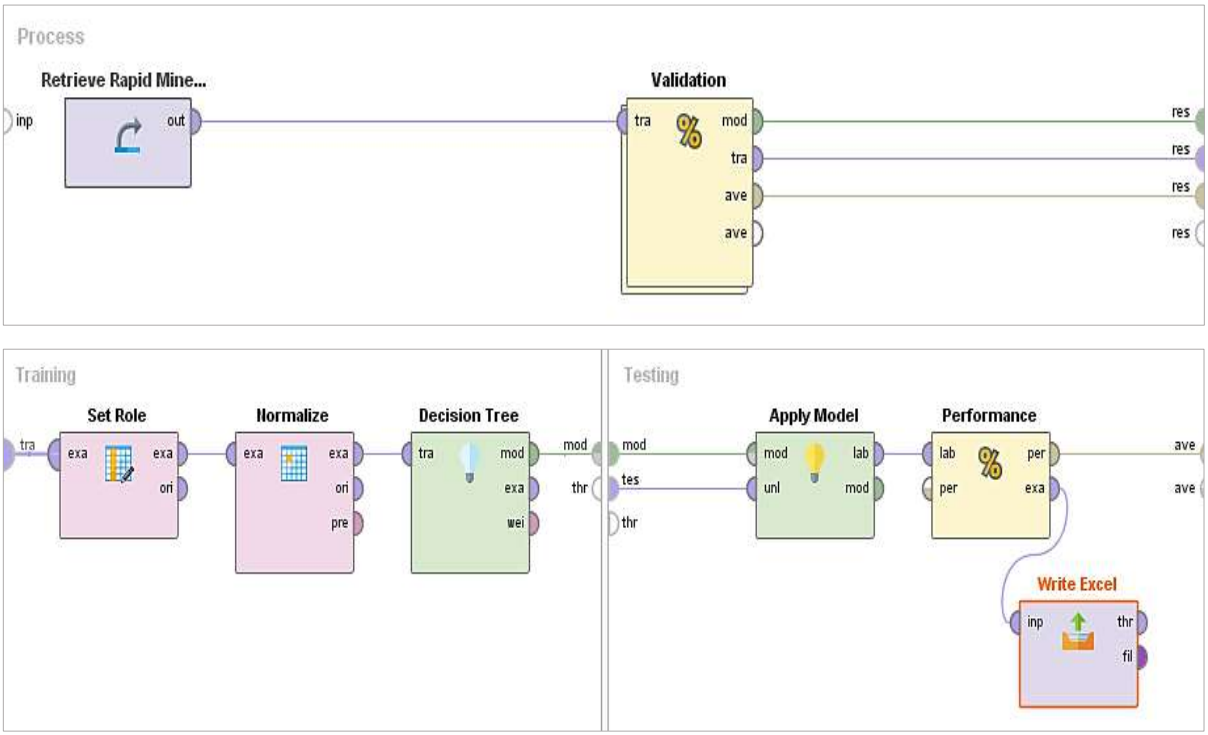


Figure 5. RapidMiner framework model

2.2.1. SVM Model Structure

The initial predictive model is constructed utilizing Support Vector Machine (SVM) tools, a supervised machine learning paradigm renowned for its ability to handle both linearly and nonlinearly separable datasets. SVM accomplishes this feat by employing a kernel function, which transforms the input space into a higher-dimensional feature space where linear separation becomes feasible. In this study, the Radial Basis Function (RBF) kernel is specifically selected due to its effectiveness in capturing intricate patterns within the dataset.

The RBF kernel calculates the similarity or proximity between two data points, facilitating the delineation of decision boundaries in the transformed feature space. A pivotal parameter governing the behavior of the RBF kernel is the kernel gamma parameter, which dictates the influence magnitude of a single training sample on the decision boundary. Smaller values of gamma signify a broader influence range, implying that each training sample has a more extensive impact, while larger gamma values restrict the influence to nearby samples, leading to a more localized decision boundary.

Table 3 outlines the SVM parameters tailored for this study. Of significant importance is the C parameter, which plays a pivotal role in balancing the trade-off between correctly classifying training examples and maximizing the margin of the decision function. A higher value of C prioritizes accurate classification of training examples, potentially leading to a narrower margin, while a lower value of C emphasizes maximizing the margin, potentially allowing for more robust generalization to unseen data. The selection of these parameters is meticulously crafted to optimize the performance of the SVM model within the context of the study's objectives. By fine-tuning the RBF kernel's gamma parameter and carefully balancing the trade-off with the C parameter, the SVM model can effectively delineate decision boundaries and make accurate predictions, thereby enhancing the study's analytical insights and predictive capabilities in engineering applications.

Table 3. Details of SVM parameters.

SVM Parameters	**M-θ Model Value**
Kernel type	Radial
C	0.0
Convergence epsilon	0.0132
L positive	1.0
L negative	1.0
Epsilon	0.0
Epsilon plus	0.0
Epsilon Minus	0.0

The RBF kernel can be written as below:

$$k(x_1, x_2) = exp(-g \parallel x_1 - x_2 \parallel^2) \quad (2)$$

Where, x_1 and x_2 are two points for function K, $\parallel x_1 - x_2 \parallel$ is the Euclidean (L_2-norm) distance between two points x_1 and x_2, and g is gamma.

2.2.2. DL model structure

The Deep Learning (DL) algorithm employed in this study represents a sophisticated approach to machine learning, characterized by its utilization of a multi-layer feed-forward Artificial Neural Network (ANN) and abstract learning methodology. Within this algorithmic framework, each iteration is executed iteratively until the output achieves a desired level of accuracy, ensuring robust predictive performance. The operational process commences with the instantiation of a local H2O cluster, configured with a single node, which serves as the foundation for the multi-layer feed-forward ANN. This neural network architecture is meticulously trained using stochastic gradient descent, leveraging the backpropagation algorithm to optimize model parameters iteratively. Notably, the training process is optimized through multi-threading, enabling asynchronous execution across individual nodes. This parallelized approach enhances computational efficiency and expedites model convergence. Furthermore, the contribution of each node to the global model is harmonized through model averaging, facilitating comprehensive ensemble learning across all compute nodes.

In the context of this specific investigation, the activation function is configured to the Rectifier Linear Unit (ReLU), a widely adopted non-linear activation function. The ReLU function is characterized by its simplicity and efficiency, as it computes the maximum between zero and the input value, effectively introducing non-linearity to the network while mitigating issues such as the vanishing gradient problem. Table 4 provides a concise summary of the key DL parameters utilized in this study, offering comprehensive insights into the configuration settings adopted for model training and optimization. The selection of these DL parameters was deliberate and tailored to the unique characteristics of the dataset under scrutiny, aiming to

strike a balance between model complexity, computational efficiency, and predictive accuracy. By leveraging the capabilities of a multi-layer feed-forward ANN trained with stochastic gradient descent and utilizing the Rectifier Linear Unit activation function, this approach seeks to harness the full potential of DL in extracting intricate patterns and making accurate predictions from the data.

Table 4. Details of DL parameters.

DL Parameters	M-θ Model Value
Activation	Rectifier
Hidden Layer Sizes	2
Numeration	50:50
Reproducible	No
Epochs	10

2.2.3. DT model structure

The DT model employed in this study utilizes a tree-structured classifier to generate predictions, leveraging the average value of the dependent variable at each leaf node. The specific configuration entails employing the "Least Square" criterion and imposing a maximal depth constraint of 10. The rationale behind selecting the "Least Square" criterion is rooted in its analogy to the least-squares optimization approach commonly utilized in linear models. This criterion guides the decision-making process within the DT by iteratively partitioning nodes into sub-nodes. At each node, the splits are strategically determined to minimize the residual sum of squares (RSS) between the observed values and the mean, with respect to the true value.

The decision to impose a maximum depth of 10 on the tree structure is significant. It serves to restrict subsequent node splitting when the tree's depth reaches an integer value of 10. This constraint is vital for preventing overfitting, ensuring that the model maintains generalizability and does not excessively tailor itself to the training data, thus enhancing its predictive performance on unseen data. Table 5 outlines the optimal DT parameters employed in this investigation, highlighting the meticulous consideration given to configuring the model for robust and accurate predictions. The chosen parameters reflect a judicious balance between model complexity and performance, tailored to the specific nuances of the dataset under scrutiny. The residual and residual sum of squares are as follows:

$$\varepsilon_i = y_i - \hat{y}_i \tag{3}$$

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i) = \varepsilon_1^2 + \varepsilon_2^2 + .. + \varepsilon_n^2 \tag{4}$$

where $\hat{y}$ is the prediction of the y_i value, n is the sample size, and ε is the difference between the actual and predicted values.

Table 5. Details of DT parameters.

DT Parameters	M-θ Model Value
Criterion	Least Square
Maximal depth	10
Pre-pruning	Applied
Minimal gain	0.1
Minimal leaf size	1

2.3. Predictive performance measurement

For every prediction model produced based on different ML algorithms, the R^2, MAE and RMSE values were measured to assess the accuracy and margin error of the predicted M-θ characteristics. These three metrics will be used to assess how well a model fits a dataset by predicting the value of the response variable using the values of the predictor variables. In the “Testing” stage of the “Split Validation” function, a “Performance”

tool is connected with an input from the "Apply Model" tool via the "Labelled Data" port and ported out from 'Performance' to the "Validation Averageable" port. The "Validation Averageable" of 'Split Validation' must be connected to the "Process Result" port to record the three-evaluation metrics. The R^2 value quantifies the level of variance that the fitted model accounted for. It is utilised frequently in the process of evaluating several models to determine which model provides the best overall fit to the data. The MAE is determined by first summing the deviations of all the predicted values from the true values of the label attribute and then dividing this total by the overall number of predictions. RMSE is the square root of the residual variance or the square root of the mean square error (RMSE). When comparing fitted values to accurate data, RMSE calculates the average squared difference between the two sets of values. In general, the lower the RMSE and MAE values, the better the model's fit to the data for predicting the experimental rotation value. In contrast, a better fit is shown by the highest or closest to $R^2 = 1$ value. Below are formulas for metrics for model evaluation:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \underline{y})^2} \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n - k}} \quad (6)$$

$$MAE = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{n} \quad (7)$$

3.0 RESULTS AND DISCUSSION

3.1. SVM performance

A comparison between the experimental rotations with the predicted ones using the SVM Radial model and modified SVM Radial model is shown in Figure 6. It can be observed from the results that the SVM Radial model (Figure 6a) overestimated the rotation value of the boltless steel connection. The discrepancy between the predicted value and the experimental value was enormous. The error can be expressed in RMSE and MAE to determine the average difference between the predicted and actual values. The RMSE and MAE of SVM Radial were found to be 1.1919 and 0.9635, respectively. In addition, the model also illustrated the low correlation between the predicted and experimental values. A modified model was proposed to decrease the SVM Radial model's prediction error. Based on the linear relationship between the predicted and experimental rotation derived from the SVM Radial scatterplot, a modification factor was calculated using Equation 8. The experimental result was compared to a newly developed set of rotation prediction values for boltless steel connections.

$$y = \frac{Predicted\ rotation - 0.2048}{12.04} \quad (8)$$

The modified SVM Radial model exhibited a notable improvement in rotation projection, as seen in Figure 6(b). The modified model effectively predicted the rotation within the experimental input range. Additionally, the modified method lowered the RMSE and MAE to 0.0541 and 0.0430, respectively. Nevertheless, the modified SVM Radial model exhibited a low positive correlation between the predicted and actual rotation values. The R^2 value for SVM Radial models was maintained at 0.2437. Furthermore, despite any negative sample data from the experimental rotation value, both SVM radial models predicted negative rotation values. This signifies the poor performance of the SVM Radial algorithm in training the sample data based on the three input variables.

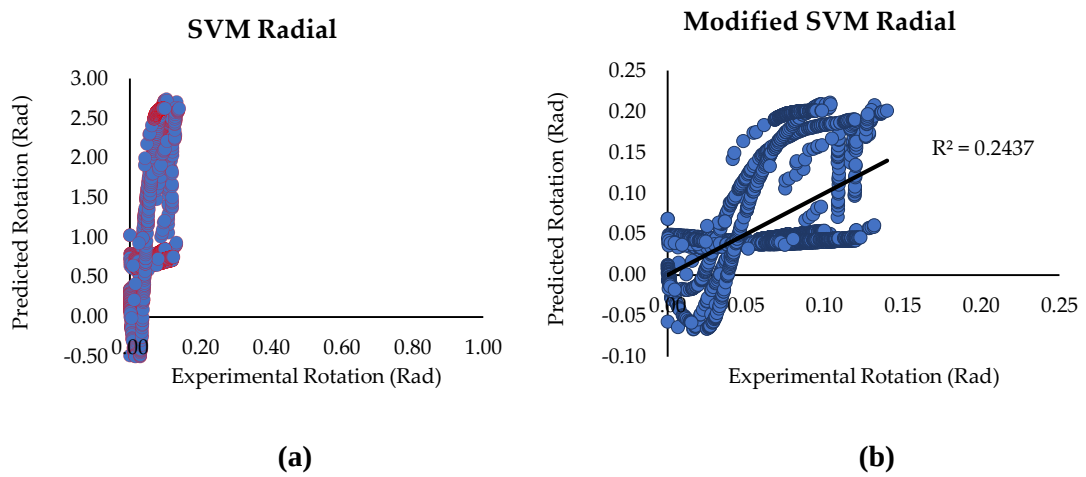


Figure 6. Comparison between experimental and predicted rotations using a) SVM Radial and b) modified SVM Radial method.

3.2. DL performance

The performance of the first DL Rectifier model to predict rotation values based on experimental rotation values is depicted in Figure 7(a). There was a significant positive correlation between anticipated and experimental rotation values, as indicated by R^2 values of 0.7915. The RMSE and MAE for the first DL Rectifier model were 0.4499 and 0.4082, respectively. Despite its substantial positive R^2 value, the initial DL Rectifier model lacked model-to-data fit due to considerable discrepancies between the predicted and actual rotation values, as seen by its high RMSE and MAE values. Nonetheless, each time the DL Rectifier method was executed, the DL Rectifier model yielded different projected rotation values. The DL Rectifier model performed slightly inferior on the second try compared to the initial effort. Figure 8(a) shows a moderately positive correlation between the predicted and actual rotation for the second attempt, which was 0.5299. Although both predicted and actual rotation values tended to increase in response to one another, the relationship was not very substantial. Compared to the initial DL Rectifier, the RMSE and MAE of this prediction model were higher at 0.5617 and 0.4918, respectively. The second DL Rectifier model was thus less precise than the first DL Rectifier model.

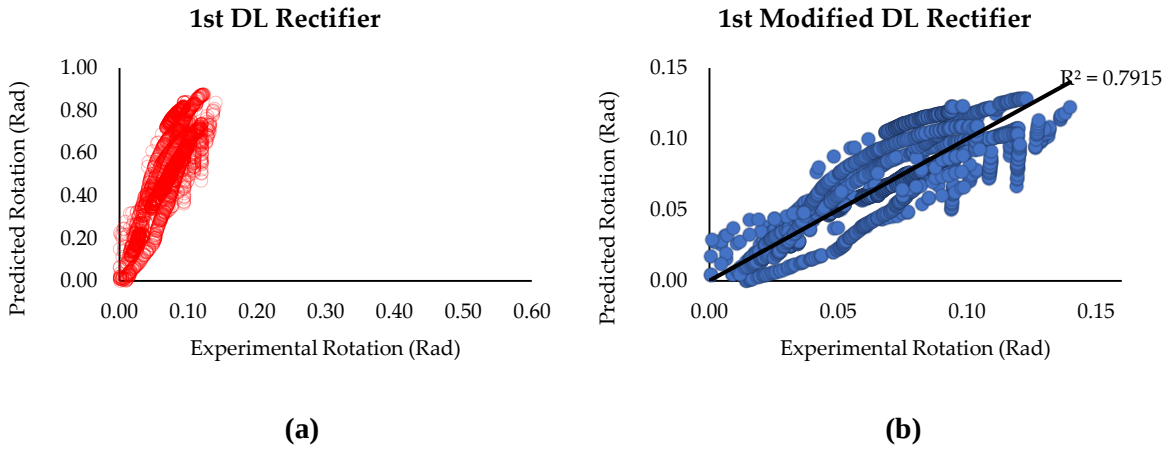
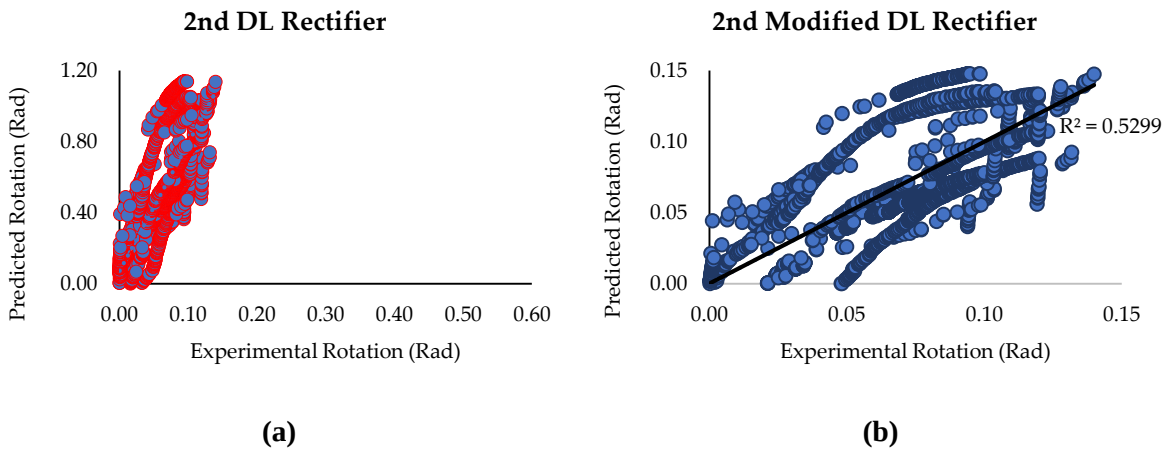


Figure 7. Comparison of rotation for a) DL 1st Rectifier and b) modified DL 1st Rectifier.



(a) **(b)**

Figure 8. Comparison of rotation for DL 2nd Rectifier and modified DL 2nd Rectifier

The results indicated that both models, on average, overestimated the rotation. Consequently, new predictions were derived for both the first and second DL Rectifiers based on the corresponding modification factors given by Equation 9 and Equation 10.

$$y = \frac{Predicted\ rotation - 0.035}{6.5404} \quad (9)$$

$$y = \frac{Predicted\ rotation - 0.0721}{0.5299} \quad (10)$$

Figure 7(b) and Figure 8(b) show the prediction performance for the first and second modified DL Rectifier models, respectively. Both modified models significantly improved in predicting the rotation within the available experimental data range for boltless steel connections. Due to modest discrepancies between the anticipated and actual values of the variable, the predicted rotation for the first modified DL Rectifier model demonstrated fewer errors. In other words, the RMSE and MAE for the first modified DL Rectifier model were smaller than the second modified DL Rectifier model. Therefore, the modified first DL Rectifier model was more capable of predicting the rotation based on experimental data.

3.3. DT performance

Figure 9 shows the comparison of rotation for DT Least Square and experimental with various splitting ratios (60:40, 70:30, 80:20). The highest accuracy of the developed predictive DT Least Square model for rotation against the actual rotation of boltless steel connection is depicted in Figure 9(c). DT Least square had the highest and closest to $R^2 = 1$ value, which was 0.8543. The prediction model exceeded a recommended rule of thumb for an acceptable R^2 of 0.75, defined as substantial. The model also has the lowest RMSE and MAE which are 04427 and 0.4065, accordingly. In addition, the model overestimated the rotation value. Equation 11 expressed the modification factor to minimize the error.

$$y = \frac{Predicted\ rotation - 0.0654}{6.1279} \quad (11)$$

As for Figure 9(d), a considerable percentage of predicted rotation points for modified DT Least Square lay near the diagonal line, indicating a high correlation between predicted and actual rotation. Most of the projected rotation from the model could be considered accurate with a slight variance compared to the experimental rotation data. Moreover, the reduction of RMSE and MAE signified the improvement in the predictions yielded from the modified model. The RMSE and MAE for modified DT Least Square were 0.0127 and 0.0094, respectively. As a result, the modified DT least squares model was found to be the most accurate prediction model for estimating the M-θ characteristics of boltless steel connections.

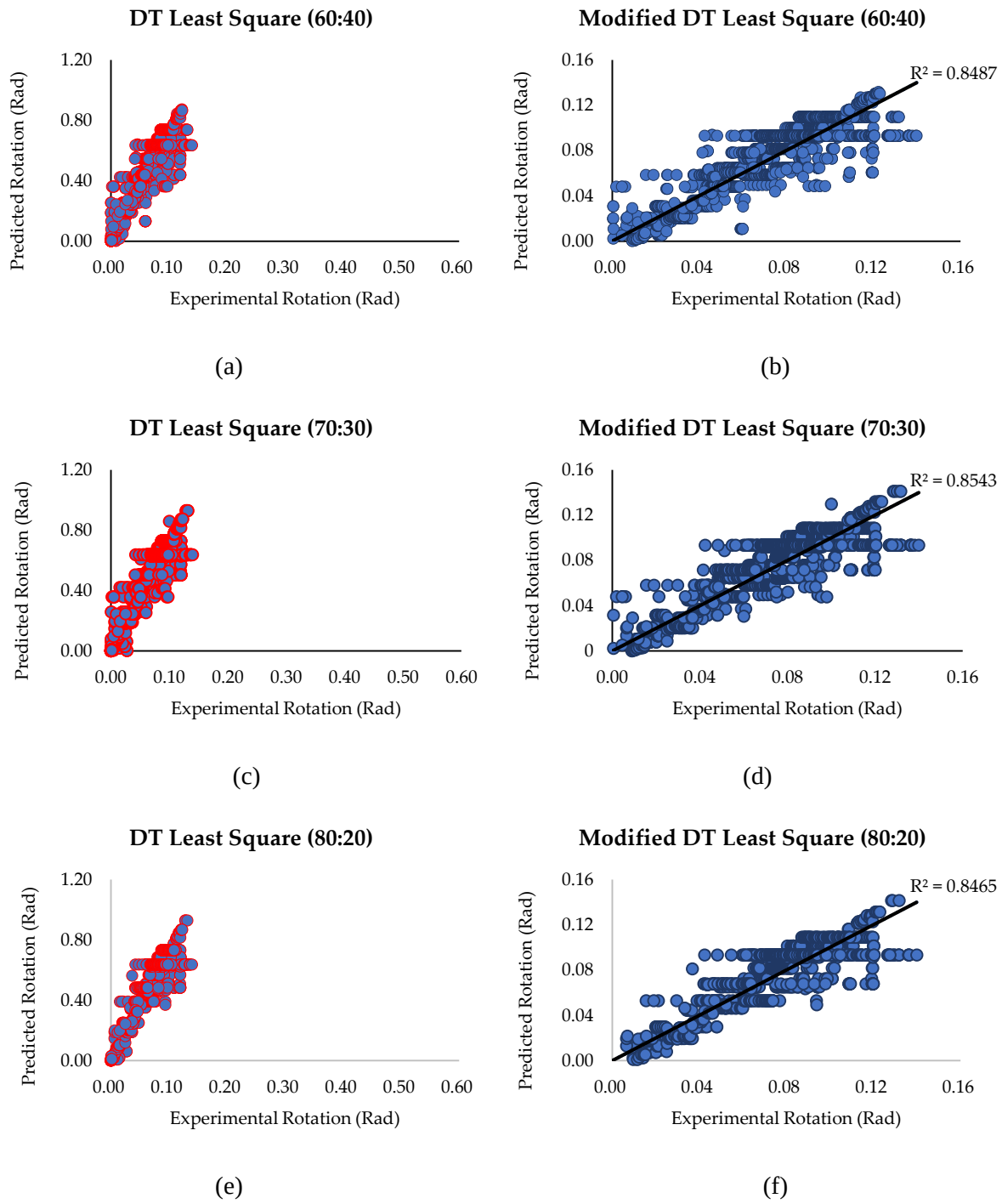


Figure 9. Comparison of rotation for DT Least Square and experimental.

3.4. Performance comparison of SVM, DL and DT

Table 6 summarizes the feasibility of different prediction models based on predictive parameters for determining the rotation of boltless steel connections based on experimental value. Among the three suggested ML algorithms, the DT Least Square model was found to be superior to the other two. In contrast, the SVM Radial model had the lowest predictive performance. The model displayed the highest RMSE and MAE values, with the lowest R^2 following adjustment. The DL algorithm was rejected because of output inconsistency. This was due to the impossibility of accurately evaluating the algorithm's capacity to forecast the M-θ properties of boltless steel connections.

Table 6. Performance comparison for different prediction models in determining M-θ characteristics.

	SVM Radial	**DT Least Square**	**DL 1st Rectifier**	**DL 2nd Rectifier**
RMSE	0.0541	0.0127	0.0158	0.0289
MAE	0.0430	0.0094	0.0118	0.0234
R^2	0.2437	0.8543	0.7915	0.5299

The proposed DT Least Square model could train the experimental data of unique M-θ characteristics of boltless steel connections. It successfully differentiated the distinct design parameters as input variables to predict the rotation as not only based on the actual M-θ value. The high accuracy of the predictions might be due to the extensive sample data provided for training purposes. Kumar and Sarathe [26] discovered that the predictive efficacy of DT is related to the amount of the sample data. The DT model also provided an easy-to-understand tree-like form of regression process. This characteristic enables the research to comprehend the splitting of nodes in order to estimate a rotation value. Initially, the DT algorithm was inadequate for applying regression and predicting continuous values. The DT algorithm tended to lose information when categorizing continuous variables into multiple categories. Despite these limitations, the proposed DT Least Square yielded the most accurate M-θ behaviour of boltless steel connections. The algorithm operated on a principle of reduction of variance based on total weighted variances between previous splits. Moreover, the DT model was prone to overfitting when the tree was designed to fit all samples in the training data set perfectly. To fit the data, even the noisy one, it generated new unneeded nodes, leading to the tree becoming too complex to interpret and resulting in a poor predicted outcome. The issue could be rectified by pruning and setting constraints on the model parameter. Two tree pruning methods exist. The first one is pre-pruning, which limits the depth of the tree to a predetermined threshold level. The second one is post-pruning, which processes the fully-grown tree after computing the cross-validation accuracy at each level.

Unlike the DT Least Square model, the SVM Radial model made a prediction solely based on the input of the M-θ behaviour concerning the experimental approach. The model did not incorporate the input variables of the eight unique design parameters, which strongly influence the overall strength and rigidity of a boltless steel connection's ability to resist failure. Consequently, the suggested SVM Radial model failed to accurately forecast the M-θ features of unique boltless steel connections. Furthermore, the SVM technique is inadequate for massive datasets compared to DT since its training complexity is so high that it becomes computationally prohibitive to train and utilize for regression purposes. In addition, SVM regression is not scalable to big datasets since each training iteration continually conducts costly kernel value calculations, which is impractical and necessitates storing the whole training dataset in memory. In addition, when the target classes overlap, the SVM prediction model performs poorly due to the presence of excessive noise in the data set. Since RBF is a low-bias kernel, this issue worsens. In this study, the experimental data sets for eight distinct structural designs are not well classified by the model. The bias for the proposed SVM model could be increased by lowering the C value to create more rigid boundaries between the variables. However, the value should not be too small to avoid overgeneralization.

In contrast to DT and SVM, the proposed DL Rectifier algorithm gave different sets of predicted rotation values for boltless steel connections for each regression. The DL Rectifier algorithm trained the data differently each time to predict the M-θ behaviour based on the input data. The discrepancy between the rotation predicted by the Rectifier model and the actual rotation was large, inconsistent, and random. Therefore, it was impossible to quantify the actual performance and capacity of the ML algorithm to predict the M-θ of boltless steel connections. Contrary to deterministic algorithms such as SVM and DT, DL was classified as a stochastic ML algorithm, in which the output was subjected to randomness and uncertainty. An inadequate sample of data from a broader population was used to fit the model, generating uncertainty. Consequently, the chosen model was occasionally capable of capturing all domain characteristics. Instead, it must generalize to unobserved conditions, therefore sacrificing some precision. Therefore, it is advised to optimize the algorithm by increasing the size of the training dataset and changing a hyper-parameter of the algorithm to reduce the output variances of the DL Rectifier model. The suggestion to increase the dataset opposed the study's objective to replace the cost limitation and hard-to-repeat experimental testing as more data were needed to feed the algorithm.

Whereas for the splitting ratio on predictive performance, Figure 9(a), (c) and (e) which exhibited 60:40, 70:30 and 80:20 shows no significant difference in rotation pattern generated based on observed experimental value. All split ratios had R^2 values greater than 0.75 with RMSE and MAE values less than 0.5, as shown in Table 7. Further modifications were carried out to fit the predicted rotation within the range of experimental input data. Equation 12 and Equation 13 express the modification factor for split ratios of 60:40 and 80:20, respectively.

Table 7. Accuracy of DT Least Square model with different split ratios to predict M-θ characteristics.

Splitting Ratio	DT Least Square	RMSE	MAE	R^2
60:40	Basic	0.4434	0.4057	0.8487
	Modified	0.0113	0.0083	0.8487
70:30	Basic	0.4427	0.4065	0.8543
	Modified	0.0127	0.0094	0.8543
80:20	Basic	0.4433	0.4073	0.8465
	Modified	0.0131	0.0087	0.8465

$$y = \frac{Predicted\ rotation - 0.0654}{6.1308} \quad (12)$$

$$y = \frac{Predicted\ rotation - 0.0691}{6.0771} \quad (13)$$

The modified model with a 60:40 split ratio performed the best based on the evaluation metrics, followed by 70:30 and 80:20. The model significantly reduced RMSE and MAE. Therefore, the modified model had the lowest RMSE and MAE while maintaining an R^2 value closer to 1. Since the modification factor was based on the linear equation between anticipated and actual rotation, the R^2 value did not change since it maintained the relationship between these two variables.

3.5. M-θ characteristics of boltless steel connections

Figure 10 (a) – (h) compares the M-θ characteristics of 8 different boltless steel connections between the modified 60:40 DT Least Square model and the experimental sample. In general, the modified 60:40 DT Least Square model generated the truest predicted rotation value compared to the other proposed models. Unlike the SVM Radial algorithm, the recommended DT Least Square algorithm had the most variable identification capacity to train the sample data based on the three design parameters: thickness column, beam, and BEC. The nodes split of the developed regression tree not only inclusive to the experimental M-θ characteristics but also undertook the properties of the bottles' steel unique designs. Regardless of the close to $R^2 = 1$ value between predicted and actual rotation described by the DT Least Square model, it had a high variance expressed by the huge value of RMSE and MAE because of the overestimated rotation output. A modified 60:40 DT Least Square model generated effectively fit the predicted M-θ characteristics to the experimental data. According to Figure 10, the negatively predicted rotation by the modified 60:40 DT model can be observed for every design case. Each of them had illustrated a clockwise rotation at BCC, which causes a beam to experience a hogging bending moment as the beam bends with a concave side upwards. However, the beam for the experimental test only suffered a positive rotation and a sagging bending moment due to the vertical load applied. In addition, there are noticeable variances between the predicted and actual M-θ behaviour for every design case, hence evidently supported based on study conducted by Bompa and Park respectively [27-28]. The average magnitude of the error expressed by RMSE for the modified 60:40 model is 0.0113. The input range determines a model's capacity to fit a dataset. Therefore, the RMSE for the modified 60:40 model is considered low, with a mean data of 0.06651 rad. Lastly, the modified 60:40 can capture the unique M-θ characteristics of boltless steel connections for multiple distinct structure designs.

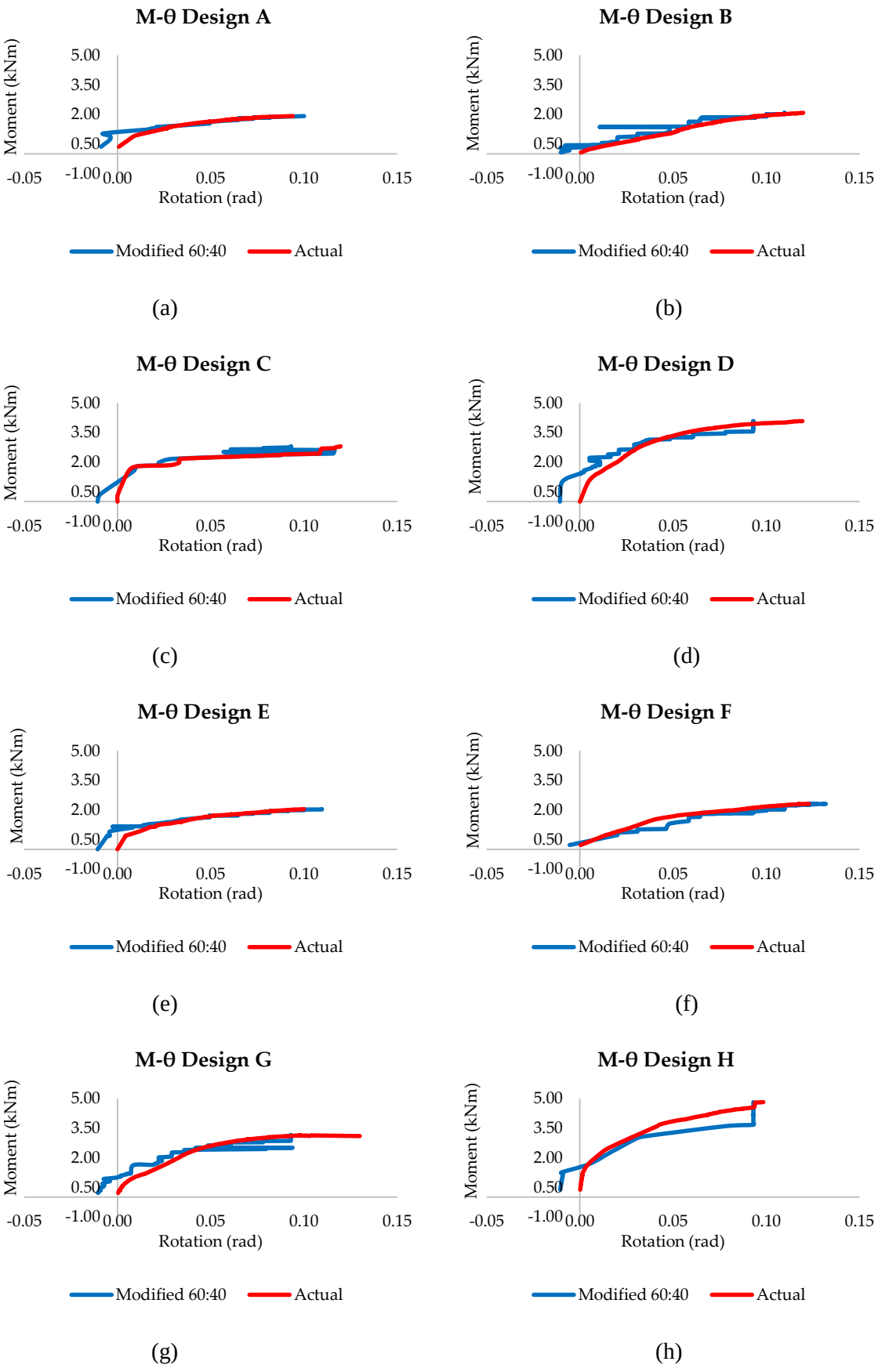


Figure 10. M-θ characteristics for different designs of boltless steel connections

4.0 CONCLUSION

This study has quantitatively demonstrated the prediction capacity of the SVM Radial, DL Rectifier and DT Least Square model to predict the M-θ characteristics of boltless steel connections based on 8020 experimental sample data. The designs of the structure heavily influenced the M-θ behaviour of boltless steel connections. This study focuses on three design inputs, namely, the thickness of the column, depth of the beam, and depth of BEC, resulting in eight unique structure design combinations of boltless steel connection. Generally, the greater the thickness of overall design inputs, the greater its rigidity and strength. The prediction performance of each model is compared based on the experimental input data. Several conclusions were made based on the three different ML algorithms to develop prediction models:

- The SVM Radial model had the lowest accuracy in predicting rotation according to the experimental value compared to DL Rectifier and DT Least Square models. As a low-bias kernel, RBF cannot differentiate input variables based on the three design parameters.
- DL Rectifier was disregarded because of inconsistency in predicting rotation. It is hard to assume the predictive capacity of the algorithm. As a stochastic ML algorithm, it requires a large training dataset and a well-defined hyper-parameter to optimise the randomness feature of the model. It is advisable to tune the hyper-parameter as a large dataset is costly and time-consuming.
- The DT Least Square outperformed other proposed models to predict the M-θ behaviour of boltless steel connections. The model shows the greatest positive correlation between the predicted and actual rotation. It is recommended to prune and set a tree depth limit to eliminate an overfitting issue often associated with the DT algorithm
- Every modified model effectively reduced the differences between the predicted and actual rotation and increased the accuracy.
- There are no significant differences between the three split ratios of 60:40, 70:30 and 80:30. Modified DT Least Square 60:40 has the lowest RMSE and MAE. Hence, it has the capacity to fit the predicted M-θ characteristics into the experimental data.

Some of the limitations were found due to the lack of variation in the boltless steel connection structural designs used for the experimental testing as input variables. The only input variables in this study are the column's thickness, beam's depth and BEC's depth. Hence the finding cannot consider the possibilities of M-θ behaviour for substantial boltless steel connections design variation in the current market. In future, more design considerations should be considered for study as input variables, for example, the type of BEC and the number of studs. The study could also include multiple steel properties available. The following limitation is unable to illustrate the performance of the ML algorithm during the training stage. Comparing performance between the training and testing with respective split ratios will help this study to identify if the proposed model is under-fitting or overfitting. Lastly, this study may also be limited by the small number of experimental sample data to develop an efficient prediction model. Specific ML algorithms, for example, DL, require a large dataset size to ensure better system performance and eliminate the inconsistency of the predicting outcomes.

ACKNOWLEDGEMENT

This research was supported by Universiti Tun Hussein Onn Malaysia (UTHM) through Tier 1 (vot Q153).

5.0 REFERENCES

Vujanac R, Miloradović N, Vulović S, Pavlović A. A comprehensive study into the boltless connections of racking systems. *Metals*. 2020 Feb 20;10(2):276.

Prabha, P., Marimuthu, V., Saravanan, M., & Jayachandran, S. A. (2010). Evaluation of connection flexibility in cold formed steel racks. *Journal of Constructional Steel Research*, 66(7), 863-872. doi: 10.1016/j.jcsr.2010.01.019

Castiglioni CA. Seismic behavior of steel storage pallet racking systems. Switzerland: *Springer International Publishing*; 2016 Feb 4.

Tsarpalis D, Vamvatsikos D, Delladonna F, Fabini M, Hermanek J, Margotan PD, Sesana S, Vantusso E, Vayas I. Macro-characteristics and taxonomy of steel racking systems for seismic vulnerability assessment. *Bulletin of Earthquake Engineering*. 2022 Mar;20(5):2695-718.

Ren C, Zhang P, Yan S, Dai L. Analysis and design of cold-formed steel storage rack uprights under localised fires. *InStructures* 2020 Oct 1 (Vol. 27, pp. 2082-2095). Elsevier.

Shah S.N., Sulong N. R., El-Shafie A. New approach for developing soft computational prediction models for moment and rotation of boltless steel connections. *Thin-Walled Structures*. 2018 Dec 1; 133:206-15.

Shah SN, Sulong NR, Shariati M, Khan R, Jumaat MZ. Behavior of steel pallet rack beam-to-column connections at elevated temperatures. *Thin-Walled Structures*. 2016 Sep 1; 106:471-83.

Zhao X, Dai L, Rasmussen KJ. Hysteretic behaviour of steel storage rack beam-to-upright boltless connections. *Journal of Constructional Steel Research*. 2018 May 1; 144:81-105.

Markazi, F. D., Beale, R. G., & Godley, M. H. R. (1997). Experimental Analysis of Semi-Rigid Boltless Connectors. *Thin-Walled Structures*, 28(1), 57-87. doi:10.1016/S0263-8231(97)00003-7

Shah, S. N. R., Sulong, N. H. R., Jumaat, M. Z., & Shariati, M. (2016). State-of-the-art review on the design and performance of steel pallet rack connections. *Engineering Failure Analysis*, 66, 240-258. doi: 10.1016/j.engfailanal.2016.04.017

Bernuzzi, C., & Castiglioni, C. A. (2001). Experimental analysis on the cyclic behaviour of beam-to-column joints in steel storage pallet racks. *Thin-Walled Structures*, 39(10), 841-859. doi:10.1016/S0263-8231(01)00034-9

Aguirre, C. (2005). Seismic behavior of rack structures. Journal of Constructional Steel Research, 61(5), 607-624. doi: 10.1016/j.jcsr.2004.10.001

Ganasan, R., Tan, C. G., Ibrahim, Z., Nazri, F. M., Sherif, M. M., & El-Shafie, A. (2021). Development of Crack Width Prediction Models for RC Beam-Column Joint Subjected to Lateral Cyclic Loading Using Machine Learning. *Appl. Sci*., 11, 7700. doi:10.3390/app11167700

Lyu, Z.-J., Wu, M., Huang, Y., Song, Y., & Cui, X. (2018). Assessment of the Dynamic Behavior of Beam-to-Column Connections in Steel Pallet Racks under Cyclic Load: Numerical Investigation. *Advances in Civil Engineering*, 2018, 12. doi:/0.1155/2018/9243216

Gurung, S. (2020). Brief study on machine learning. doi:10.13140/RG.2.2.10196.12169

Boswell, D. (2002). Introduction to Support Vector Machines. Retrieved from https://home.work.caltech.edu/~boswell/IntroToSVM.pdf

Ahmadi, A., Han, D., & Moridi, A. (2015). Input selection for long-lead precipitation prediction using large-scale climate variables: a case study. *Journal of Hydroinformatics*, 17(1), 114-129.

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. doi:10.1007/s42979-021-00592-x

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electron Markets*, 31, 685-695. doi:10.1007/s12525-021-00475-2

Mathew, A., & Arul, A. (2021). Deep Learning Techniques: An Overview. doi:10.1007/978-981-15-3383-9_54

Feizollah, A., Abdullah, N. A., Anuar, N. B., & Hanif, H. (2019). Halal Products on Twitter: Data Extraction and Sentiment Analysis Using Stack of Deep Learning Algorithms. *IEEE Access*, 7, 1-1. doi:10.1109/ACCESS.2019.2923275

Urban, C. J., & Gates, K. M. (2021). Deep Learning: A primer for psychologists. Psychological Methods, 26(6), 743-737. doi:10.1037/met0000374

Rokach, L., & Maimon, O. (2005). Decision Trees. The Data Mining and Knowledge Discovery Handbook, 6. doi:10.1007/0-387-25465-X_9

Kalita, J. (n.d.). Decisions and Regression Tree Learning. Retrieved from http://www.cs.uccs.edu/~jkalita/work/cs586/2013/DecisionTrees.pdf

Sá, J. A. S. d., Rocha, B. R. P. d., Almeida, A., & Mota, M. A. S. (2016). Lightning Forecast Using Data Mining Techniques on Hourly Evolution of the Convective Available Potential Energy. ChemBioChem, 1-5. Retrieved from 10.21528/CBIC2011-27.1

Kumar, M., & Sarathe, A. K. (2016). Efficient Prediction of Surface Roughness Using Decision Tree. *American Journal of Engineering Research*, 5(12), 228-243.

Bompa DV, Elghazouli AY. Inelastic cyclic behaviour of RC members incorporating threaded reinforcement couplers. *Engineering Structures*. 2019 Feb 1; 180:468-83.

Park AY, Wang YC. Joint rotation behaviour of bolted endplate connections to flowdrilled RHS columns. *Tubular Structures XIII, The University of Hong Kong*, ISBN 978-0-415-58473. 2010 Nov 12; 9:37-4.