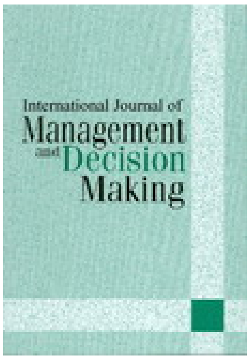


International Journal of Management and Decision Making

ISSN online: 1741-5187 - ISSN print: 1462-4621
https://www.inderscience.com/ijmdm

A hybrid approach to evaluate employee performance using MCDA and artificial neural networks

Malik Haddad, David A. Sanders

DOI: 10.1504/IJMDM.2024.10059635

Article History:

Received:	28 March 2022
Last revised:	29 August 2022
Accepted:	29 August 2022
Published online:	04 December 2023

A hybrid approach to evaluate employee performance using MCDA and artificial neural networks

Malik Haddad*

Northeastern University – London,
UK
Email: malik.haddad@nulondon.ac.uk
*Corresponding author

David A. Sanders

School of Mechanical and Design Engineering,
University of Portsmouth,
Portsmouth, UK
Email: david.sanders@port.ac.uk

Abstract: A new hybrid approach for employee performance evaluation based on multiple criteria decision analysis (MCDA) and artificial neural network (ANN) is presented. This is the first time this type of ANNs has been used for this application. A deep ANN is created. A MCDA method used randomly generated sets for training and testing the ANN. The network provided 93.63% training accuracy and 91.91% testing accuracy when tested against the training and testing sets respectively. The new approach could be transformed into a generic employee evaluation tool suitable to accommodate any number of employees and evaluation criteria using transfer-learning. A real-life employee evaluation problem is used as an example. Six employees and six evaluation criteria are considered. The new approach successfully identified the employee most eligible for promotion and ranked the other employees according to their performance.

Keywords: artificial neural networks; employee evaluation; employee performance; MCDA; transfer-learning.

Reference to this paper should be made as follows: Haddad, M. and Sanders, D.A. (2024) 'A hybrid approach to evaluate employee performance using MCDA and artificial neural networks', *Int. J. Management and Decision Making*, Vol. 23, No. 1, pp.58–76.

Biographical notes: Malik Haddad is an Assistant Professor in Data and Computer Science at Northeastern University, London. He received his BSc in Electronic Engineering in 2006 from Princess Sumaya University for Technology. He went on to postgraduate study and was awarded a distinction in MSc in Electrical and Computer Engineering in 2007 from the New York Institute of Technology. Finally, he completed his PhD in Intelligent Decision Making Applied to Management and Engineering in 2019 from the University of Portsmouth. His current research interests include intelligent decision making and multiple criteria decision making, assistive technologies, obstacle avoidance, machine learning, neural networks, artificial intelligence, intelligent systems and human machine interfaces. He has a CEng MIET, MIEEE, MIMechE and PMP – PMI.

David A. Sanders has an MBE TD VR CEng and received two first degrees, a BA in Social Science and Mathematics in 1985 (OU), and a BSc in Engineering in 1984 (Portsmouth, UK). He went on to postgraduate study and was awarded a DMS in Management Studies and then an MBA in Business Administration and Change Management in 2001 (OU). He completed his PhD in Automatic Path Planning with constraints in 1990 (CNAA, UK). Since 2019, he has been a Professor of Systems Engineering and reader in Knowledge Engineering in the Faculty of Technology at the University of Portsmouth. He is an author of three books, more than 250 articles, and several patents and inventions. His current and previous research interests include assistive technologies, sensor integration, robotics, AI, tele-operation, powered wheelchairs and obstacle avoidance. He is a Fellow IET, Fellow IMechE, and Fellow HEA.

1 Introduction

A new hybrid approach is presented to assist with employee performance evaluation based on Neural Networks and Multiple Criteria Decision Analysis (MCDA). The new approach was trained and tested, then applied to an employee performance evaluation problem presented in Haddad et al. (2019a). Results from training and testing the new approach showed that it behaved satisfactorily and achieved 93.63% training accuracy and 91.91% testing accuracy. Outcomes from applying the new approach to the real-world example were compared to the results presented in Haddad et al. (2019a) and benefits of the new approach are provided.

Research has investigated links between Human Resource (HR) systems and performance evaluation for more than 20 years in the private sector (Pak et al., 2020) as well as in the public sector (Battaglio, 2020). The Increased attention to Human Resource Management (HRM) has led to many theories and methods for workforce performance evaluation (Keegan et al., 2018). These theories and methods improved organisations' operation and efficiency (Alvehus, 2018; Op-de-Beeck et al., 2018; Jardioui et al., 2019), lowered employee turnover (Kim et al., 2018) and enhanced employee knowledge, abilities and skills and prospect to contribute (Hooi, 2019). But different evaluation methods often generated different outcomes (Haddad et al., 2019a) which could reduce motivational forces.

Mumtaz et al. (2022) mentioned three factors to affect employee turnover: organisational fairness, managerial fairness and power distance. They claimed that the employee turnover rate could be minimised by improving bonds between employees and their superiors, and receiving robust support and guidance from their superiors, management and organisation. This paper describes a new, simple and fair approach to employee performance evaluation. The new approach uses MCDA methods to evaluate a total score for the randomly simulated employees used in testing and training the Artificial Neural Network (ANN). This is the first time this type of ANNs have been applied in this way for this type of problems. An example of the application of the new approach is presented. It was applied to a real-life problem for US coast guard officer evaluation considered in Haddad et al. (2019a).

2 Employee performance evaluation

Ho and Kuvaas (2020) considered Human Resource Management (HRM) systems as a combined set of HRM routines that could improve an organisation's economic performance. Employee performance evaluation is often considered one of the earliest HRM processes (Poovathingal and Kumar (2018). A fair model for employee performance evaluation can be essential for organisation success (Haddad et al., 2019a). Many researchers studied the link between HRM practices and organisation performance and discovered a positive interaction between HRM routines and performance (Chadwick and Li., 2018). Moreover, researchers found that with a moderate application, HRM practices could be positively associated with many factors that might improve an organisation's efficiency for example staff well-being, employee commitment and job satisfaction. López-Cotarelo (2018) asserted that there was a positive interaction between the application of HRM routines and workforce performance. Ogbonnaya and Messersmith (2019) claimed that HRM practices could improve performance by bringing employees' interests closer to organisation goals. Lin and Kellough (2019) considered employee evaluation as a tool to improve employee incentives and help employees tune the way they work to better accomplish organisational goals.

Glaister et al. (2018) stated that HRM concentrated on the organisation's workforce. Almarzooqi et al. (2019) studied the effect of sustainable HRM practices on organisation success and employee performance and concluded that there was a direct positive relationship between these factors. Kim et al. (2018) encouraged building good social relationships and shared communication between employees and supervisors to reduce employee turnover. Huselid (2018) encouraged HR professionals and supervisors to create a holistic understanding of the workforce contribution toward organisational success and to reflect this understanding on the workforce performance evaluation and metrics. López-Cotarelo (2018) stated that one of the factors that differentiated HRM practices from traditional personnel management was the transfer of workforce management duties from personnel experts to line managers which increased the line managers' accountability for the outcomes of their decisions. Op-de-Beeck et al. (2018) highlighted the importance of supervisors' role in HRM implementation. Moreover, López-Cotarelo (2018) mentioned that due to line managers' daily interaction with employees, production process and consumers, they could optimise both the HRM decision-making process and business performance.

Employee performance evaluation could often be seen as an action concerning the relationship between the professionals and their managers (Lidinska and Jablonsky, 2018). Employee performance evaluation is usually seen as a multifaceted activity that considered various features and multiple evaluation criteria. Employees often participated in multiple projects and their total performance was a combination of their performance in those projects (Lidinska and Jablonsky, 2018). Participatory practices showed different effects on employee performance (Berdicchia and Masino, 2019). Studies conducted by Schuh et al. (2018) showed that employees with high-quality "leader-member exchange" relationships often received higher evaluation performance ratings from their supervisors when engaged in innovative work activities.

Anjomshoae et al. (2019) proposed a new employee performance evaluation approach based on Analytical Hierarchy Process (AHP) and a balance scoreboard method. Gabelica et al. (2016) suggested a performance evaluation approach based on two criteria: task analysis and goal achievement. Lidinska and Jablonsky (2018) presented the main fields for a performance management context: evaluation criteria; evaluation occurrence; criteria importance; and assessment approach. Levenson (2018) introduced an approach that could improve an organisation's effectiveness and strategy execution by conducting workforce analytics through the application of system diagnostics and added two additional steps that preceded the analysis: 'competitive advantage analytics' and 'enterprise analytics'. Jardioui et al. (2019) suggested dynamic performance evaluation systems that could quickly cope with internal and external changes and provide a swift and reliable response to these changes. Lopes et al. (2018) presented a predictive approach for lawyers' annual performance ranking based on an ANN, their approach provided successful outcomes and achieved 71% accuracy. The work presented in this paper used a deep long-short-term-memory (LSTM) ANN to create a generic and dynamic employee performance evaluation system that could accommodate any number of employees and evaluation criteria, cope with internal and external changes and provide a quick, reliable and accurate outcome.

A systematic review conducted by Vrontis et al. (2021) showed that Artificial Intelligence (AI), robotics and other advanced techniques were increasingly used in HRM. These techniques were employed in employee recruitment, training and performance evaluation. Laitinen and Kadak (2019) directly linked the management of performance evaluation with organisational performance. A fair and transparent employee performance evaluation model that used advanced mathematical approaches could help with employee satisfaction and organisational success. Using advanced mathematical techniques such as AI and ANN could improve organisation performance, HRM services and practices (Pan et al., 2021) and provide a more straightforward and fairer performance evaluation. That could in turn enhance employee satisfaction.

Vrbova and Mullerova (2021) claimed that the quality of decisions was highly dependent on available data and strong arguments. Pelissari et al. (2021) identified three different types of uncertainty in decision problems. Employee performance evaluation problems were often susceptible to all of these types of uncertainty. Moreover, employee performance evaluation often involved many factors and had key effects on organisation efficiency. It is a complicated and multifaceted problem. ANN could quickly produce reliable outcomes to assist in the performance evaluation when a large number of employees and evaluation criteria are considered.

3 LSTM neural network

ANNs are algorithms often applied to represent complex and nonlinear mathematical functions that are similar to employee performance evaluation functions.

A neuron is the main element of an ANN. Neurons executed complex nonlinear mathematical calculations. Neurons are simple elements taking a vector of real-valued number as an input and producing a single real-valued number as an output (Staudemeyer and Morris, 2019). Neurones have an internal input called bias. The neurone takes a vector of real-valued input values, all of which are weighted by a multiplier. The weights are initialised at the beginning of the training, as the training progress, the neurone adjusts these weights based on training data. The Neurone sums all weighted input values and fires if the resultant value is above a predefined threshold (Staudemeyer & Morris, 2019).

ANNs are inspired by biological learning systems and aim to model their functions. Biological learning systems are complex networks of interconnected neurons (Staudemeyer & Morris, 2019). ANNs are usually assembled by putting together multiple neurons to form a layer, and combining several layers to construct an ANN. A deep ANN is often composed of an input layer, an output layer and several hidden layers in between. A fully connected ANN consists of a series of fully connected layers where each neuron in one layer is connected to all neurons in the next layer."

A learning algorithm is used for training a network on a set of samples (for example employees) during training. To train a Network, forward and backward propagation is conducted. In forward propagation, input data move forward from the input layer, through the hidden layers to the output layer to provide an output. The learning algorithm assigns weights to each input of the neurons. In backward propagation, the weights assigned to the input of the neurons are fine-tuned to minimise the difference between the true outcome and the ANN outcome. This is accomplished by tuning the weights by a small value called the "learning rate" and then processing all example employees. As the training set goes through multiple nonlinear layers with each layer identifying specific features of the set. Processing data through the layers, the network would be able to recognise the appropriate identifiers for accurately classifying the data into appropriate classes. Processing all example employees in a set, tuning the weights and minimising variance is called an epoch. The set considered in this paper consisted of 10,000 example employees, and the networks considered in this paper completed 100 -150 epochs before testing.

Different structures of neural networks were often popular for specific sorts of applications. A deep LSTM ANN is considered in this paper. LSTM ANN was first presented in 1997 by Hochreiter and Schmidhuber (1997). Since then, the structure of LSTM ANN had been simplified and its function had been improved to achieve better proficiency and precision (Greff et al., 2016).

LSTM Neural Networks are one of the most powerful classifiers known (Staudemeyer & Morris, 2019). LSTM are a type of Recurrent Neural Networks (RNNs) (Williams & Zipser, 1989; Werbos, 1990). RNNs are dynamic systems, they have connections between higher and lower layer neurones and optional self-feedback connections. These connections enable RNNs to build a memory of time series events by creating an internal state for every step of the classification process and enable data flow from earlier events to current steps (Staudemeyer & Morris, 2019). RNNs range from partly connected to fully connected networks. The network used in this approach is a fully connected network.

The growth of statistical prediction and modelling led researchers to adopt more complex algorithms and approaches, such as ANNs, to tackle difficulties in identifying patterns and predictions (Olden and Jackson, 2002). ANNs performed well compared to other traditional approaches (Rafiee Parsa et al., 2021). ANNs are getting greater attention as an effective, flexible and reliable modelling technique for predicting patterns (Graves et al., 2008). LSTM has been effectively applied to most fields of science such as text completion, vehicle trajectory, handwriting recognition and pattern recognition (Dai et al., 2019; Graves et al., 2008). Employee performance evaluation problems can be similar.

4 The new approach applied to employee performance evaluation

In the work presented in this paper, input data were fed through multiple nonlinear layers of a fully connected ANN, with individual layers identifying different patterns of data. When data were processed throughout layers, ANN could map the input sequence to a specific output identifier to classify the data into suitable classes. Due to their excellent predictive and classification powers, ANN could generate reliable and suitable results for employee performance evaluation problems.

The new approach could be used for different employee evaluation problems that consider different numbers of employees and evaluation criteria. To apply the new approach to the real-life employee evaluation problem presented in Section 5, the approach considered six employees and six evaluation criteria.

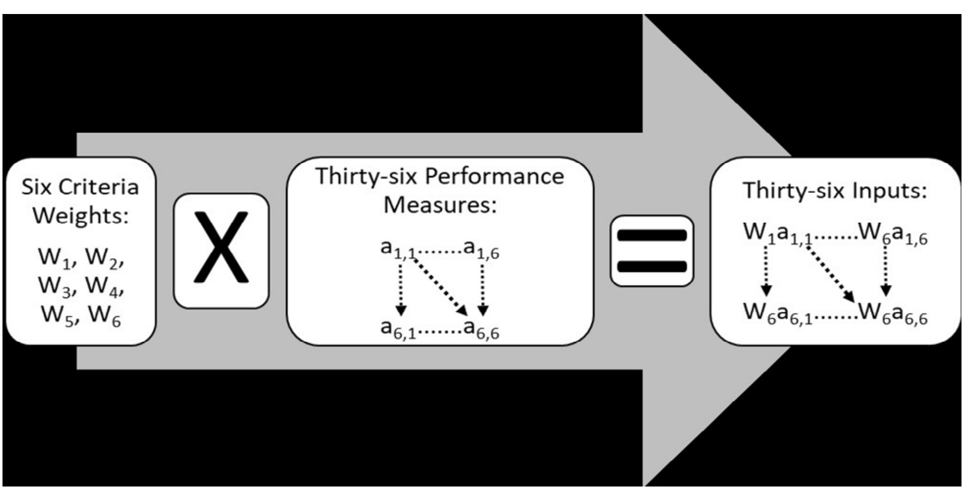
Randomly generated employee performance measures and criteria weights were created using MS Excel. Ten thousand sets were created, each set consisted of 36 performance measures and 6 criteria weights. Sets and the total scores are considered as inputs to train and test a LSTM ANN. Weighted Sum Model (WSM) was used to calculate the total score of the employees:

$$P_y = \sum_{x=1}^{n} W_x . a_{xy} \tag{1}$$

where P_y represented the total score of employee y, W_x represented the weight of criterion x, and a_{xy} represented the performance measure of employee y with regard to criterion x. The most suitable answer was usually the answer that had the highest value of P_y. The total scores were considered to identify the most suitable employee for promotion and rank the employees.

The ANN used considered thirty-six inputs and six outputs. The performance of each employee with regard to each criterion was scaled to create the 36 inputs to the ANN. 6 officers were rated against the 6 criteria to produce 36 inputs. Inputs were the 6 weights multiplied by the 36 performances as shown in Figure 1.

Figure 1 Multiplying criteria weights by performance measures to create inputs for the neural network

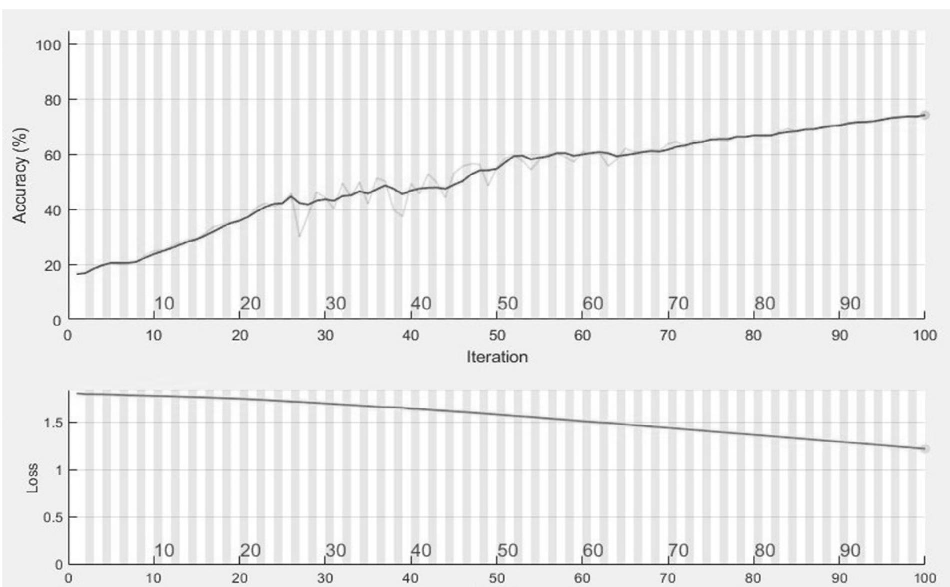


The LSTM neural network used was the same as the network used in Haddad and Sanders (2020).

MATLAB platform was used to create the layers, the structure of the layers and to set options for training the ANN. ANN was built with 36 input units, 100 hidden units in the BILSTM Layer, and 6 classification classes. The same training algorithm and training options used in Haddad and Sanders (2020) were used.

A (10,000x36) matrix was used to train and test the ANN. The matrix was split into a 3:7 ratio for testing and training respectively. Figure 2 shows ANN training. ANN training accuracy is shown in the upper section of Figure 2 and the ANN loss is shown in the lower section of Figure 2. ANN training accuracy provided an understanding of the performance of the ANN and measured how well the ANN performed over each epoch. It assessed the ANN result with respect to the correct result. The ANN training loss signified an overall number of mistakes made for each example. Training loss represented how weakly the ANN performed after each epoch. To optimise ANN performance, ANN training accuracy had to be maximised and ANN training loss had to be minimised.

Figure 2 Network training progress with 36 inputs and 6 output classes, by the end of 100 epochs, network training accuracy had increased to 74.2% (above)

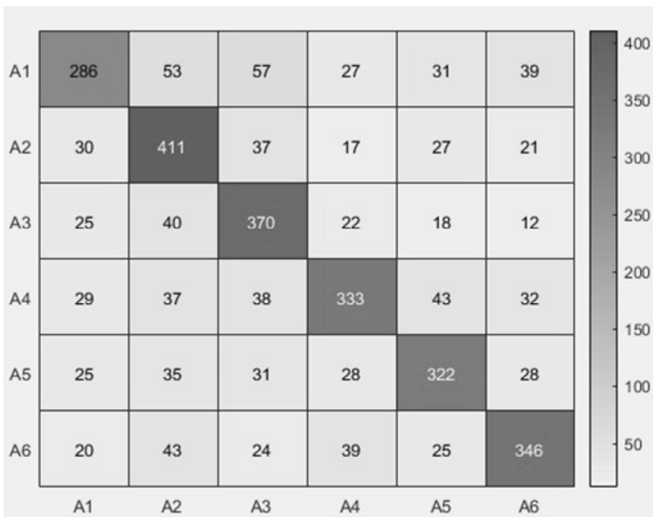


As training progressed, training loss decreased and training accuracy increased. When training finished, the ANN training accuracy was higher than 74%. Network testing accuracy was 69.4% when tested using the testing set.

Figure 3 shows a confusion matrix generated from testing the ANN. A confusion matrix shows the network outcome (Vertical Axis) vs the correct result (Horizontal axis). The ANN correct outcomes are shown diagonally (darker cells). The diagonal darker cells represented the true positives i.e., how many times the model correctly predicted the right outcome, while the non-diagonal cells represented the false positives i.e., how many times the model incorrectly predicted the outcome. The network correctly predicted A_1

(Officer U) 286 times but incorrectly predicted A_1 (Officer U) 53 times as A_2 (Officer V), 57 times as A_3 (Officer W), 27 times as A_4 (Officer X), 31 times as A_5 (Officer Y) and 39 times as A_6 (Officer Z). The network correctly predicted A_2 (Officer V) 411 times but incorrectly predicted A_2 (Officer V) 30 times as A_1 (Officer U), 37 times as A_3 (Officer W), 17 times as A_4 (Officer X), 27 times as A_5 (Officer Y) and 21 times as A_6 (Officer Z). The network correctly predicted A_3 (Officer W) 370 times, A_4 (Officer X) 333 times, A_5 (Officer Y) 322 times and A_6 (Officer Z) 346 times.

Figure 3 Confusion chart used to assess network accuracy after completing 100 epochs considering 36 inputs and 6 output classes



The ANN produced reasonable accuracy. Other values for training options were used in an attempt to improve accuracy. A compromise between training accuracy and time was made.

Figure 4 shows network training progress. As ANN training progressed training accuracy increased, as shown in the upper section of Figure 4 and ANN loss decreased, as shown in the lower section of Figure 4. After 150 epochs, the training accuracy reached 84.73%. That was an improvement.

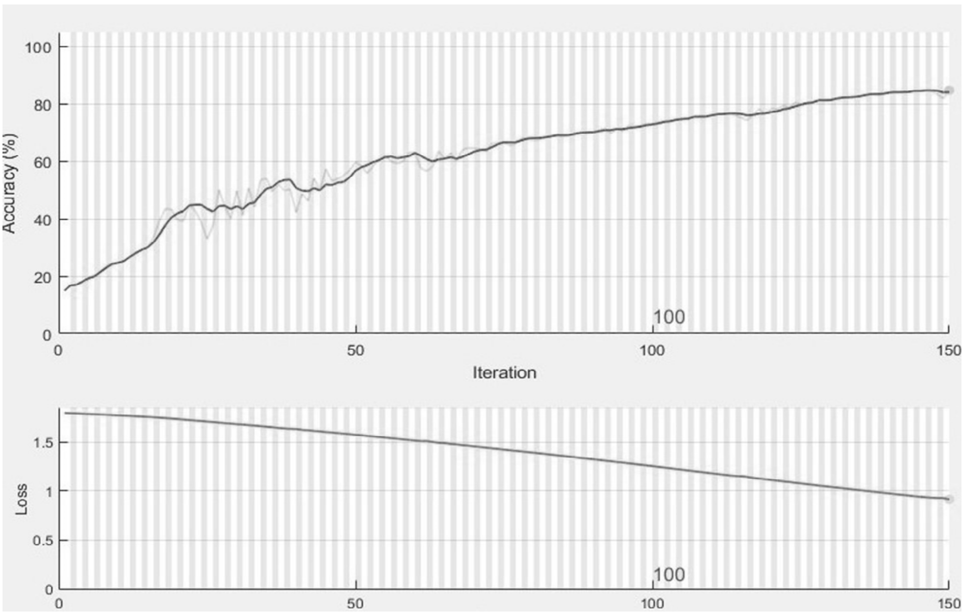
Figure 4 confirmed that as training advanced, training loss decreased and training accuracy increased. After 150 epochs the ANN training accuracy reached 84.73%.

The network achieved 80.21% accuracy when tested. Increasing the number of epochs to 150 improved network accuracy.

Figure 5 shows the confusion matrix generated from testing the network considering thirty-six inputs and six outputs after processing 150 epochs. The ANN correct outcomes are shown diagonally (darker cells). The diagonal darker cells represented the true positives i.e., how many times the model correctly predicted the right outcome, while the non-diagonal cells represented the false positives i.e., how many times the model incorrectly predicted the outcome. The network correctly predicted A_1 (Officer U) 333 times but incorrectly predicted A_1 (Officer U) 56 times as A_2 (Officer V), 36 times as A_3 (Officer W), 22 times as A_4 (Officer X), 29 times as A_5 (Officer Y) and 17 times as A_6 (Officer Z). The network correctly predicted A_2 (Officer V) 506 times, A_3 (Officer W)

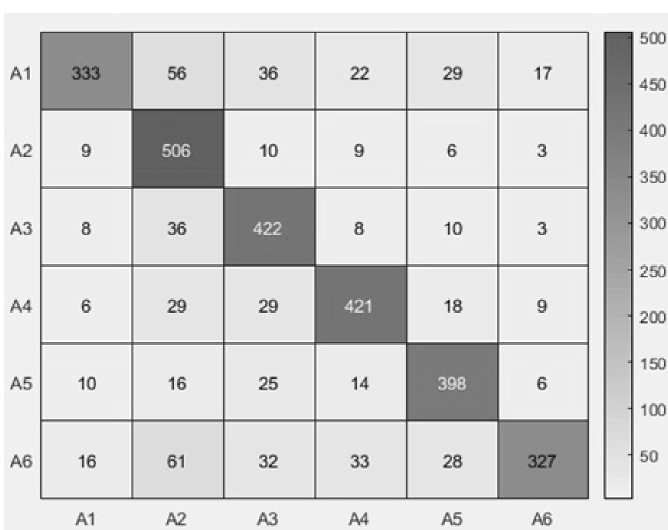
422 times, A_4 (Officer X) 421 times, A_5 (Officer Y) 398 times and A_6 (Officer Z) 327 times. The confusion matrix in Figure 5 shows a larger number of correct outcomes (darker cells) with respect to the number of correct outcomes presented in the confusion matrix in Figure 3.

Figure 4 Network training progress with 36 inputs and 6 outputs



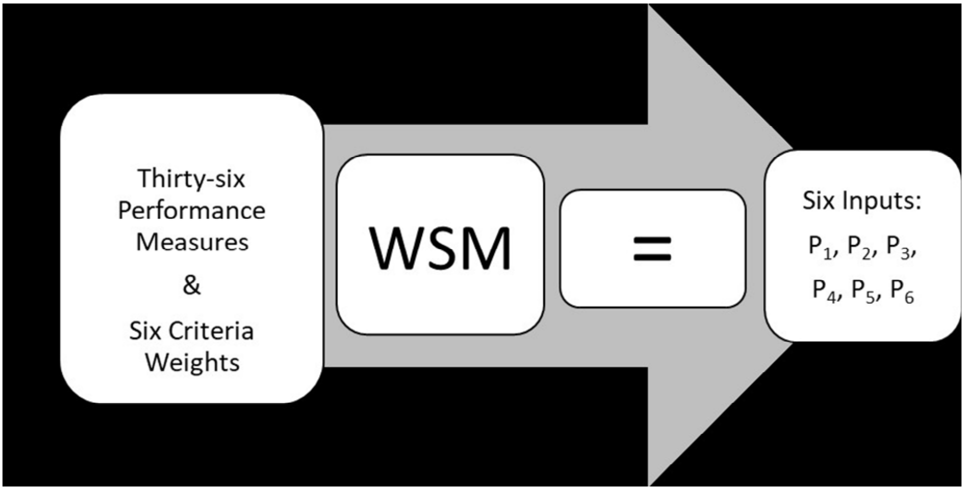
Note: After completing 150 epochs, network training accuracy had increased to more than 80% (above) and network loss had decreased (below).

Figure 5 Confusion chart used to assess network accuracy after completing 150 epochs and considering 36 inputs and 6 output



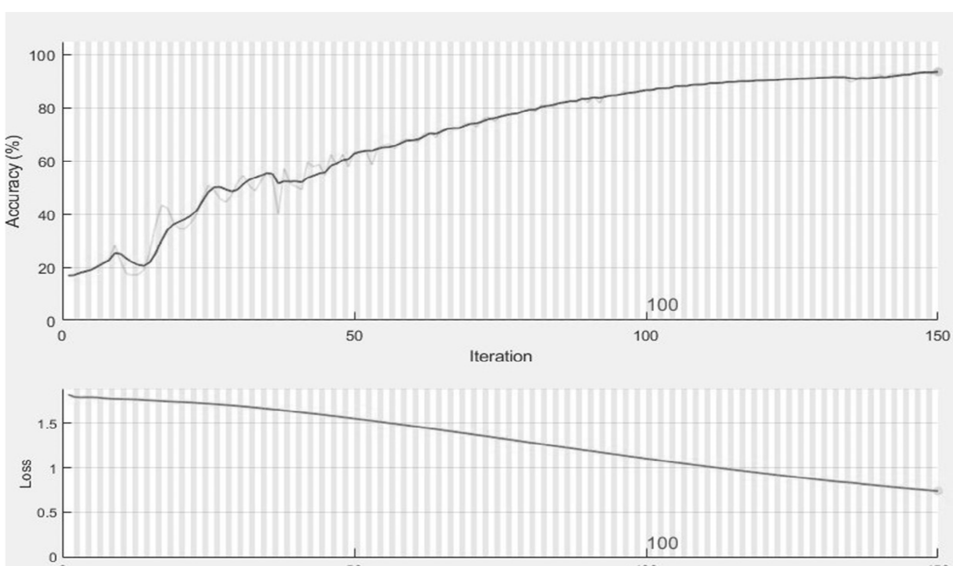
Inputs to the ANN were reformed in an attempt to improve accuracy further. WSM was applied to calculate the total score of the officers using equation 1. Inputs to the LSTM network could then be reduced to six as shown in Figure 6.

Figure 6 Reducing the number of inputs to the neural network using WSM



Equation (1) was used to calculate the total scores of the officers. Six total scores were calculated. The total scores were used as inputs to the ANN.

Figure 7 Network training progress considering 6 inputs and 6 output classes, after completing 150 epochs, the network training accuracy had increased to more than 90% (above) and the network loss had decreased (below)



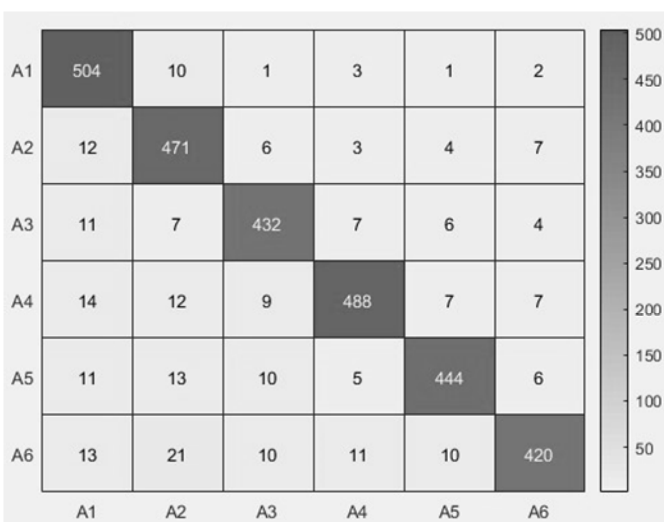
The same network structure and training algorithm were used in this architecture with 0.001 as a learning rate and 150 epochs were considered. The same testing and training sets were considered.

Figure 7 shows the ANN training progress considering 6 inputs and 6 outputs. Network performance improved as training progressed. The ANN training accuracy increased shown in the upper section of Figure 7. ANN loss decreased shown in the lower section of Figure 7. After 150 epochs the ANN training accuracy reached 93.63%. Reducing the number of inputs improved the network training accuracy and reduced the time required to train the network.

ANN accuracy reached 91.91% when tested. That was a significant improvement from the previous two cases.

Figure 8 presents a confusion matrix generated from testing the ANN with six inputs and six outputs after processing 150 epochs. The ANN correct outcomes are presented diagonally (darker cells). The diagonal darker cells represented the true positives i.e., how many times the model correctly predicted the right outcome, while the non-diagonal cells represented the false positives i.e., how many times the model incorrectly predicted the outcome. The ANN correctly predicted A1 (Officer A) 504 times but incorrectly predicted A_1 (Officer A) 10 times as A_2 (Officer B), once as A_3 (Officer C), 3 times as A_4 (Officer D), once as A_5 (Officer E) and twice as A_6 (Officer F). The network correctly predicted A_2 (Officer B) 471 times, A_3 (Officer C) 432 times, A_4 (Officer D) 488 times, A_5 (Officer E) 444 times and A_6 (Officer F) 420 times. Comparing results from Figures 3, 5 and 8 confirms that the model generated a larger number of true positives and a smaller number of false positives. The ANN generated higher accuracy when the number of inputs was decreased to 6, and the number of epochs was increased to 150.

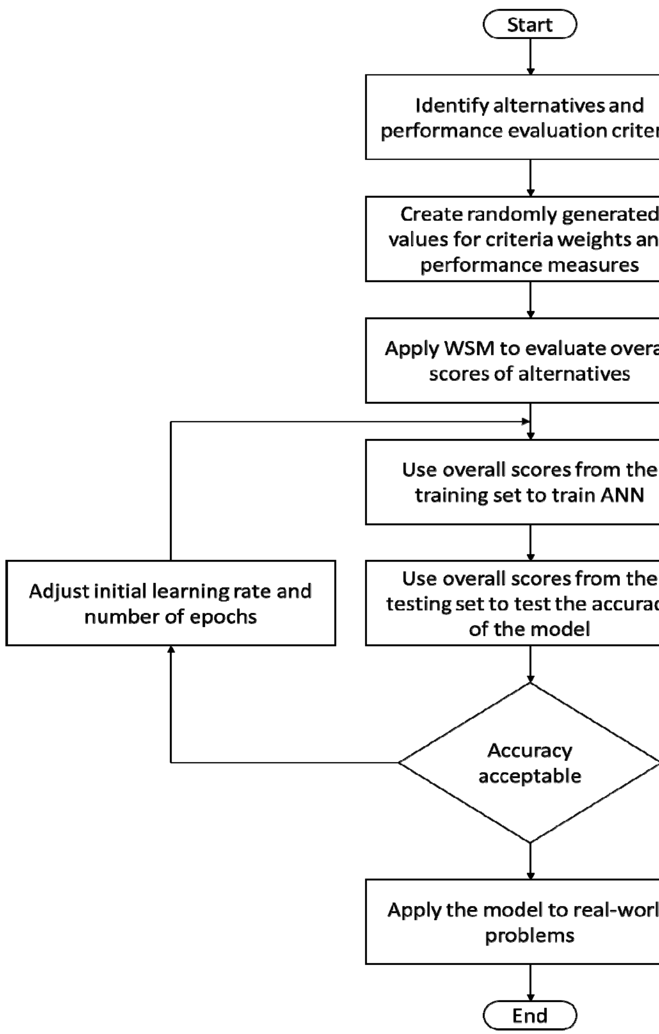
Figure 8 Confusion chart used to assess network accuracy after completing 150 epochs and considering 6 inputs and 6 output



The network required less time to process 150 epochs and produced higher accuracy when the number of inputs was decreased. The learning rate and the number of epochs were set to 0.001 and 150 respectively.

Figure 9 shows a flowchart of the new approach. The flowchart shows a step-by-step operation of the new approach. The first step in the new approach was to identify the alternatives and evaluation criteria. The second step was to create randomly generated values for the criteria weights and performance, these values would be used for training and testing the neural network used in the new approach. The third step applied the WSM method to calculate the overall scores of all alternatives with respect to all criteria. The fourth step used the overall scores from the training set to train the neural network, in the fifth step, the overall scores in the testing set were used to assess the accuracy of the new approach. If the decision makers accepted the accuracy of the model, the new approach would be applied to the real-world problem, if the accuracy was not acceptable to the decision makers, then the model parameters would be adjusted, and the training and testing process would be repeated until the accuracy of the model is accepted by the decision makers.

Figure 9 Flowchart of the new approach



5 The new approach applied to a specific employee performance evaluation problem

The new approach was applied to an employee performance evaluation problem considered in Haddad et al. (2019a). This problem assessed the eligibility of U.S. Coast Guard officers for promotion. A set of 6 criteria were considered, and a set of 6 anonymous employees were evaluated. The 6 evaluation criteria were the same as the evaluation criteria considered in Haddad et al. (2019a). Criteria names and weights are shown in Table 1.

Table 1 Decision matrix for the employee evaluation considered in Haddad et al. (2019a)

Criteria and criteria weights \ *Alternatives*	A_1 *Officer A*	A_2 *Officer B*	A_3 *Officer C*	A_4 *Officer D*	A_5 *Officer E*	A_6 *Officer F*
C_1: Performance of duties = 0.296	0.152	0.181	0.172	0.170	0.172	0.153
C_2: Interpersonal relations = 0.254	0.172	0.150	0.176	0.161	0.150	0.191
C_3: Leadership skills = 0.159	0.193	0.156	0.186	0.166	0.162	0.137
C_4: Communication skills = 0.125	0.183	0.173	0.150	0.174	0.150	0.170
C_5: Personal qualities = 0.084	0.196	0.170	0.161	0.162	0.157	0.155
C_6: Representing the coast guard = 0.082	0.170	0.203	0.142	0.175	0.164	0.145

Officer B had the best performance with respect to Performance of Duties and was given the highest score of 0.181, followed by Officers C and E. Both officers performed similarly and were given the same score of 0.172, followed by Officers D and F. Their scores were 0.170 and 0.153 respectively. Officer A had the weakest performance with respect to Performance of Duties and was given the lowest score of 0.152. Officer F had the best performance with respect to Interpersonal Relations and was given the highest score of 0.191, followed by Officers C, A and D. Their scores were 0.176, 0.172 and 0.161 respectively. Officers B and E had the weakest performance with respect to Interpersonal Relations. Their score was the lowest. They were given the same score of 0.150. Officer A had the best performance with respect to Leadership Skills and was given the highest score of 0.193, followed by Officers C, D, E and B. Their scores were 0.186, 0.166, 0.162 and 0.156 respectively. Officer F had the weakest performance with respect to Leadership Skills and was given the lowest score of 0.137. Officer A had the best performance with respect to Communication Skills and was given the highest score of 0.183, followed by Officers D, B and F. Their scores were 0.174, 0.173 and 0.170 respectively. Officers C and E had the weakest performance with respect to Communication Skills. Their score was the lowest. They were given the same score of 0.150. Officer A had the best performance with respect to Personal Qualities and was given the highest score of 0.196 followed by Officers B, D, C and E. Their scores were 0.170, 0.162, 0.161 and 0.157 respectively. Officer F had the weakest performance with respect to Personal Qualities and was given the lowest score of 0.155. Officer B had the best performance with respect to Representing the Coast Guard and was given the highest score of 0.203, followed by Officers D, A, E and F. Their scores were 0.175, 0.170, 0.164 and 0.145 respectively. Officer C had the weakest performance with respect to Representing the Coast Guard and was given the lowest score of 0.142.

Haddad et al. (2019a) applied two popular Multiple Criteria Decision Analysis (MCDA) methods to an employee performance evaluation problem. The two methods were the Preference Ranking Organization METHod for Enrichment Evaluations II, (PROMETHEE II) and AHP.

Results presented in Haddad et al. (2019a) showed that PROMETHEE II generated the following ranking of officers:

1 Officer C.

2 Officer B.

3 Officer A.

4 Officer D.

5 Officer F.

6 Officer E.

AHP provided a slightly different ranking of officers:

1 Officer A.

2 (Officer B = Officer C).

3 Officer D.

4 Officer F.

5 Officer E.

The approach described in this paper was applied to the same employee performance evaluation problem considered in (Haddad et al., 2019a).

The trained and tested ANN was used to identify the officer most eligible for promotion and to rank the officers based on their performance. WSM was used to model the opinion of the HR team and to create inputs for the ANN. Six inputs were considered.

The new approach identified (Officer A) as the most eligible officer for promotion and generated the following ranking of officers:

1 Officer A.

2 Officer C.

3 Officer B.

4 Officer E.

5 Officer F.

6 Officer D.

The total score of the officers was:

- Officer A = 0.1695.

- Officer B = 0.166.
- Officer C = 0.1689.
- Officer D = 0.1645.
- Officer E = 0.1658.
- Officer F = 0.1652.

6 Discussion and results

A new hybrid employee evaluation approach based on ANN and MCDA was presented. The new approach was trained and tested using randomly generated training and testing sets. The new approach showed a satisfactory outcome and achieved 91.91% when tested. The trained and tested approach was applied to a real-life employee performance evaluation problem considered in Haddad et al. (2019a), it successfully identified Officer A as the most eligible officer for promotion. This result was the same as the result provided by AHP in Haddad et al. (2019a) and by the rules used to train the ANN. The new approach provided a different ranking than the rankings produced by AHP and PROMETHEE II. (Haddad et al., 2019a).

MATLAB was used to apply the new approach, MATLAB is often seen as simple, mathematically inexpensive and did not need professional experience.

A first attempt to create the new approach considered 36 inputs and six outputs. The performance of each employee with regard to each criterion was scaled to create the 36 inputs to the ANN but that did not yield good accuracy.

A second attempt was conducted to improve approach accuracy. The second attempt considered the same number of outputs but with other values for the training option and yielded better accuracy than the first attempt.

A third attempt was conducted to increase accuracy and considered the same number of outputs but considered reducing the number of inputs to the ANN from 36 to 6 by applying WSM to calculate the overall scores of the alternatives with respect to criteria. The overall weights of the alternatives were used as inputs to the ANN. The attempt used the same training options considered in attempt 2 and yielded better accuracy than the previous 2 attempts. Attempt 3 yielded a training accuracy of 93.63% and 91.91% accuracy testing accuracy. Reducing the number of inputs improved the network training accuracy and reduced the time required to train the network. The third attempt was a successful attempt that yielded the required accuracy.

Interaction between criteria was not considered in this approach. No indifference, preference or veto thresholds were considered and the alternative that had a higher performance measure on a criterion was preferred to other alternatives with respect to that criterion.

The new approach could be generalised using Transfer-Learning (Côté-Allard, 2019) to accommodate for a larger number of employees and evaluation criteria. It could also be generalised to tackle any ranking or choice problems for example corporate relocation, strategic marketing and supplier selection problems.

7 Conclusions and future work

A new approach for employee performance evaluation based on an ANN and MCDA is presented. The new approach was successfully applied to US Coast Guard officers' performance evaluation problem. This is the first time this type of ANNs was used for this type of application. The approach mixed ANN with MCDA concepts and provided a simple, reliable and efficient employee evaluation system by using advantages from both concepts.

A real-life HRM example considered in Haddad et al. (2019a) was used to demonstrate the simplicity and efficiency of the new approach. The results produced were validated against the results in Haddad et al. (2019a) and the results generated from the rules used to train the ANN. The new approach provided satisfactory outcomes and was straightforward, transparent and more efficient than the decision making methods used in Haddad et al. (2019a). Moreover, the new approach did not need experience or understanding of MCDA. The results of the new approach could be considered to identify the most eligible employee for promotion or to rank the employees based on their performance.

The authors are currently using Python Programming Language to create a similar approach and are creating different types of ANN using other programming languages and applying the new approaches to other examples. A comparison between the different types of ANNs will be conducted.

The new approach could be transformed into a more generalised employee evaluation tool using transfer-learning that could accommodate any number of employees and evaluation criteria.

Transfer learning will be applied to transform the new approach into a more generic approach that could accommodate a larger number of alternatives (Officers) and a larger number of performance evaluation criteria.

Future work will apply the approach to a larger number of alternatives (officers) and performance evaluation criteria.

Future work will also consider using deeper ANN as well as more theoretically advanced MCDM methods like AHP and PROMETHEE.

References

Almarzooqi, A.H., Khan, M., and Khalid, K. (2019) 'The role of sustainable HRM in sustaining positive organizational outcomes: An interactional framework', *International Journal of Productivity and Performance Management,* Vol. 68 No. 7, pp. 1272-1292.

Alvehus, J. (2018) 'Conflicting logics? The role of HRM in a professional service firm', *Human Resource Management Journal,* Vol. 28, No. 1, pp. 31–44.

Anjomshoae, A. et al (2019) 'An integrated AHP-based scheme for performance measurement in humanitarian supply chains', *International Journal of Productivity and Performance Management,* Vol. 68 No. 5, pp. 938-957.

Battaglio, P. (2020) 'The Future of Public Human Resource Management', *Public Personnel Management,* Vol. 49 No. 4, pp. 499-502.

Berdicchia, D. and Masino, G. (2019) 'The Ambivalent Effects of Participation on Performance and Job Stressors: The Role of Job Crafting and Autonomy', *Human Performance,* Vol. 32 No. 5, pp. 220-241.

Chadwick, C. and Li, P. (2018) 'HR systems, HR departments, and perceived establishment labor productivity', *Human Resource Management,* Vol. 57, No. 1, pp. 1415–1428.

Côté-Allard, U.et al (2019) 'Deep learning for electromyographic hand gesture signal classification using transfer learning', *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* Vol. 27, No. 4, pp. 760-771.

Dai, S. et al (2019) 'Modeling vehicle interactions via modified LSTM models for trajectory prediction', *IEEE Access,* Vol. 7, No. 1, pp. 38287-38296.

Gabelica, C. et al (2016) 'Establishing team knowledge coordination from a learning perspective', *Human Performance,* Vol. 29, No. 1, pp. 33-53.

Glaister, A.J. et al (2018), 'HRM and performance—The role of talent management as a transmission mechanism in an emerging market context', *Human Resource Management Journal,* Vol. 28, No. 1, pp. 148–166.

Glatte, T. (2015) 'Location Strategies: Methods and their methodological limitations', *Journal of Engineering Design and Technology,* Vol. 13, No. 3, pp. 435 – 462.

Graves, A. et al (2008) 'A novel connectionist system for unconstrained handwriting recognition', *IEEE Transactions on Pattern Analysis and machine intelligence,* Vol 31, No. 5, pp. 855–868.

Greff, K. et al (2016) 'LSTM: A search space odyssey', *IEEE transactions on neural networks and learning systems,* Vol. 28, No. 10, pp. 2222-2232.

Haddad, M. and Sanders, D. (2019) 'Selecting a best compromise direction for a powered wheelchair using PROMETHEE', *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* Vol. 27, No. 2, pp. 228-235.

Haddad, M. et al (2020) 'Artificial Neural Network Approach for Business Decision Making applied to a Corporate Relocation Problem', *Archives of Business Research*, Vol. 8, No. 6, pp. 180–195.

Haddad, M. et al (2020) 'Selecting a discrete multiple criteria decision making method for Boeing to rank four global market regions', *Transportation Research Part A: Policy and Practice,* Vol. 134, No. 1, pp. 1-15.

Haddad, M. et al (2019a) 'Selecting a robust decision making method to evaluate employee performance', *International Journal of Management and Decision Making,* Vol. 18, No. 4, pp. 333 – 351.

Haddad, M. et al (2019b) 'Selecting a Discrete Multiple Criteria Decision Making Method to decide on a Corporate Relocation', *Archives of Business Research,* Vol. 7 No. 5, pp. 48-67.

Ho, H. and Kuvaas, B. (2020) 'Human resource management systems, employee well-being, and firm performance from the mutual gains and critical perspectives: The well-being paradox', *Human Resource Management,* Vol. 59, No. 1, pp. 235–253.

Hochreiter, S. and Schmidhuber, J. (1997) 'Long short-term memory', *Neural Computation,* Vol. 9, No. 8, pp.1735-1780.

Hooi, L.W. (2019) 'Firm performance: is organizational learning capability the magic wand?', *International Journal of Productivity and Performance Management,* Vol. 68, No. 8, pp. 1411-1433.

Hopfield, J.J. (1988) "Artificial neural networks", *IEEE Circuits and Devices Magazine*, Vol. 4, No. 5, pp. 3-10.

Huselid, M.A. (2018) 'The science and practice of workforce analytics: Introduction to the HRM special issue', *Human Resource Management,* Vol. 57, No. 1, pp. 679–684.

Jardioui, M. et al (2019) 'How organizational culture influences performance measurement systems in SMEs', *International Journal of Productivity and Performance Management,* Vol. 69, No. 2, pp. 217-235.

Keegan, A. et al (2018) 'Organizing the HRM function: Responses to paradoxes, variety, and dynamism', *Human Resource Management,* Vol. 57, No. 1, pp. 111-1126.

Kim, S. et al (2018) 'The "HR–line-connecting HRM system" and its effects on employee turnover', *Human Resource Management,* Vol. 57, No. 1, pp. 1219–1231.

Laitinen, E.K., and Kadak, T. (2019) 'Explaining success of performance management systems: The role of chain of key factors', *International Journal of Productivity and Performance Management,* Vol. 68, No. 2, pp. 362-388.

Levenson, A. (2018) 'Using workforce analytics to improve strategy execution', *Human Resource Management,* Vol. 57, No. 1, pp. 685–700.

Lidinska, L. and Jablonsky, J. (2018) 'AHP model for performance evaluation of employees in a Czech management consulting company', *Central European Journal of Operations Research,* Vol. 26, No. 1, pp. 239 - 258.

Lin, Y.C. and Kellough, J.E. (2019) 'Performance appraisal problems in the public sector: Examining supervisors' perceptions', *Public Personnel Management,* Vol. 48, No. 2, pp. 179-202.

Lopes, S.A. et al (2018) 'Can artificial neural networks predict lawyers' performance rankings?', *International Journal of Productivity and Performance Management,* Vol. 67, No. 9, pp. 1940-1958.

López-Cotarelo, J. (2018) 'Line managers and HRM: A managerial discretion perspective', *Human Resource Management Journal,* Vol. 28, No. 1, pp. 255–271.

Mumtaz, R. et al (2022) 'Investigating managerial and fairness practices on employee turnover intentions through the mediation of affiliation quality between organisation and employee. A comprehensive study of the metropolitan society of Malaysia', *International Journal of Management and Decision Making*, Vol. 21, No. 1, pp.1-27.

Ogbonnaya, C. and Messersmith, J. (2019) 'Employee performance, well-being, and differential effects of human resource management subdimensions: Mutual gains or conflicting outcomes?', *Human Resource Management Journal,* Vol. 29, No. 1, pp. 509–526.

Olden, J.D. and Jackson, D.A. (2002) 'Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks', *Ecological modelling,* Vol. 154, No. (1-2), pp. 135-150.

Op-de-Beeck, S. et al (2018) 'Explaining effective HRM implementation: A middle versus first-line management perspective', *Public Personnel Management,* Vol. 47, No. 2, pp. 144-174.

Pak, K. et al (2020) 'The influence of human resource practices on perceived work ability and the preferred retirement age: A latent growth modelling approach', *Human Resource Management Journa*l, Vol. 33, No. 1, pp. 311–325.

Pan, Y. et al (2021) 'The adoption of artificial intelligence in employee recruitment: The influence of contextual factors', *The International Journal of Human Resource Management,* pp. 1-23. DOI: 10.1080/09585192.2021.1879206.

Pelissari, R. et al (2021) 'Techniques to model uncertain input data of multi-criteria decision-making problems: a literature review', *International Transactions in Operational Research,* Vol. 28, No. 2, pp. 523-559.

Poovathingal, J.S. and Kumar, D.V. (2018) 'Quantifying the contribution of competencies toward performance', *International Journal of Productivity and Performance Management,* Vol. 67, No. 4, pp. 693-716.

Rafiee Parsa, N. et al (2021) 'A hybrid neural network approach to minimize total completion time on a single batch processing machine', *International Transactions in Operational Research,* Vol. 28, No. 5, pp. 2867-2899.

Williams, R.J. and Zipser, D. (1989) "A learning algorithm for continually running fully recurrent neural networks", *Neural Computation,* Vol. 1, No. 2, pp. 270-280.

Schuh, S.C. et al (2018) 'Are you really doing good things in your boss's eyes? Interactive effects of employee innovative work behavior and leader–member exchange on supervisory performance ratings', *Human Resource Management,* Vol. 57, No. 1, pp. 397–409.

Staudemeyer, R.C. and Morris, E.R. (2019) "Understanding LSTM--a tutorial into long short-term memory recurrent neural networks", *arXiv preprint arXiv:1909.09586*.

Vrbová, L. and Müllerová, B. (2021) "Informal reasoning fallacies: answered and unanswered questions from a decision-making perspective", *International Journal of Management and Decision Making,* Vol. 20, No. 1, pp.112-122.

Vrontis, D. et al (2021) 'Artificial intelligence, robotics, advanced technologies and human resource management: a systematic review', *The International Journal of Human Resource Management,* pp. 1-30. DOI: 10.1080/09585192.2020.1871398.

Werbos, P. J. (1990). Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, Vol. 78, No. 10, pp. 1550-1560.