Vocational Education Skill Assessment and Intelligent Assistance: A Study on the Application of Machine Learning Algorithms in the Assessment of Vocational Information Literacy Teaching Ability

Jiali Dai[1] and Hanifah Jambari[2]

Abstract

Investigating vocational educators' knowledge-based teaching skills across China's Vocational Education (VE) institutions, this research focuses on the practical use of Machine Learning (ML) algorithms. Instructors' efficacy must be evaluated, and this work addresses the gap. VE performs an essential role in connecting learning abilities with the demands of industry. The investigation plans on developing an adaptable, subjective assessment technique that extends within the boundaries of conventional subjective evaluation methods using modern ML techniques such as Support Vector Machines (SVM), Decision Trees (DT), and Neural Networks (NN). Each ML model's accuracy, reliability, and feasibility have been determined using data collected from 120 vocational educators encompassing various fields and regions. Researchers predict that our findings will provide perspective on how to improve vocational education settings' teaching methods and governance.

Keywords: *Vocational Education, Machine Learning, Educational Practices, Policy-Making, Precision.*

INTRODUCTION

An effective technique for measuring teaching skills has to be found in the rapidly evolving field of Vocational Education (VE), which is driven by rapid innovations in technology and constant financial requirements. Attention to the quality of education is at the highest level ever throughout China due to the vital role of VE in supporting both economic and industrial development. Knowledge of technology is an essential skill in today's digital age, and the present research addresses a significant gap by analyzing the skills of vocational educators. In the nation of China, VE plays a vital role since it bridges the gaps between traditional educational systems and the demands of current job seekers. The teaching methods used by instructors in these educational settings are required to adapt to the developing manufacturing sectors [4]. Developing a robust Information Literacy Teaching Ability (ILTA) is essential for students to perform well in jobs and has been recognized as a key aspect of VE.

There is an extreme shortage of reliable, scalable techniques to evaluate teachers' achievement in the field of knowledge-based education, considering that this is a field that has been recognized to be essential. Many present methods for evaluation utilize subjective assessments or indirect evaluations, which may not represent the most appropriate method to assess a teacher's effectiveness or what outcomes their instruction achieves for the pupils they teach. Educator appraisal and Information Literacy Teaching Ability (ILTA) have been the subject of many recent studies in HE, but the individual setting of VE has received relatively little study attention. In addition, nearly all of these research projects use qualitative methods which are either not accurate sufficient or aren't simple to scale. Current approaches to assessing educators in vocational environments can frequently be discontinuous and attempt to make use of technological advancements that could improve accuracy and data analysis. There is a shortage of adaptability and success in informing educational policy and practice due to the reality that numerous research efforts fail to adequately account for the different and ever-changing contexts in which VE plays action.

The purpose of this study is to provide a novel approach for evaluating the ILTA of China's vocational teachers in several settings using Machine Learning (ML) algorithms. A more objective, reliable, and scalable review

[1] School of Education, Universiti Teknologi Malaysia, Johor, 81310, Malaysia, Lucheng School, Wenzhou Polytechnic, Wenzhou, Zhejiang, 325000, China E-mail: jiali202024@outlook.com (Corresponding Author)

[2] School of Education, Universiti Teknologi Malaysia, Johor, 81310, Malaysia E-mail: Hanifah Jambari: hanifah-j@utm.my

method is the objective of this research, which uses modern computational methods like Neural Networks (NN), Decision Trees (DT), and Support Vector Machines (SVM). The findings from this study will help address gaps in the theoretical understanding of what makes for effective teaching methods and will also provide practical suggestions for how to improve training for teachers and curriculum.

The research paper has been organized in the following order: the summary of the literature is given in part 2, the approach is shown in section 3, the results and analysis will be displayed in section 4, and the work is concluded in section 5.

METHODOLOGY

Study Setting and Sample Population

The current study presented an in-depth review of the educational system in the region by studying an ample number of VE institutions across China. Also, carefully selected vocational instructors from ten distinct VE institutions. These educators teach in many vocational and technical professional fields of study, encompassing culinary arts and automotive technology. The following institutions have been recognized for their superior VE and were selected for their comprehensive academic services, which serve an enormous number of students from various socioeconomic groups and regions.

These higher education institutions represent variations of China's academic environment, extending from big urban centers like Shanghai and Beijing to rural and smaller cities. The success and generalizability of ML models designed for assessing ILTA rely on this variability faculty from those educational institutions enrolled. A stratified sampling method was used to select participants so that the sample would approximate the wide range of academic backgrounds, professional seniorities, and experiences identified in Chinese VE.

Table 1 illustrates the demographic data of the participants, which was developed with care to ensure diversity. The ages of the participants varied from 25 to 60 years, and there was a balanced distribution of genders. Their different educational backgrounds encompass teaching methods, technical fields, and VE. In order to ensure that the findings are accurate and feasible in practical scenarios, this socioeconomic variation is essential for the research's objective of testing the results of ML algorithms on an accurate group of vocational teachers across China.

Table 1: Demographic details

Demographic	Description
Total Participants	120
Gender Distribution	58 Male, 62 Female
Age Range	25-60 years
Age Group Distribution	25-35 years (35), 36-45 years (45), 46-60 years (40)
Location Distribution	Urban Centers (36), Suburban Areas (48), Rural Areas (36)
Institutions	10 Vocational Colleges across China
Educational Background	VE (38), Technical Fields (52), Pedagogy (30)
Professional Seniority	Less than 5 years (34), 5-15 years (52), More than 15 years (34)

Participants scored their responses on a Likert scale from 1 (Strongly Disagree) to 7 (Strongly Agree), enabling an in-depth assessment of vocational educators' ILTA as educators. The questionnaire comprised 13 questions designed to gauge various aspects of teaching competency and attitudes toward ILTA in vocational settings. Additionally, two factual questions were included to gather data on teaching experience and educational background, which are crucial control variables that do not utilize the Likert scale format.

The variable description of the questionnaire is presented below:

Independent Variables (Professional Development & Technology Use)

Professional Development: I seek professional development opportunities to enhance my ILTA.

Technology Integration: I am proficient in integrating technology to facilitate ILTA among students.

Resource Utilization: When teaching students in ILTA, I make effective use of several tools, such as databases, online resources, and multimedia.

Feedback Utilization: In order to enhance my ILTA instruction, I frequently employ student feedback and ideas.

Peer Collaboration: In order to come up with novel ideas for promoting ILTA, I regularly collaborate with other educators.

Dependent Variables (Teaching Effectiveness)

Self-Efficacy in Teaching: I am confident in my ability to teach ILTA effectively.

Curriculum Integration: Including the topic of ILTA into my curriculum is automatic.

Student Outcome Satisfaction: The ILTA that my students have developed fulfilled my criteria.

Critical Thinking Emphasis: I emphasize the development of critical thinking skills through ILTA education.

Instructional Challenges: I effectively overcame challenges encountered in ILTA.

Mediating Variables (Engagement Strategies)

Student Engagement I regularly implement strategies that actively engage students in ILTA.

Adaptability to Change: I quickly adapt my teaching methods to incorporate new ILTA standards and technologies.

Cultural Sensitivity: I tailor my ILTA teaching to accommodate cultural diversity in my classroom.

Control Variables (Experience & Background)

Teaching Experience: How many years have you been teaching in VE? (Please specify the number of years.)

Educational Background: What is your highest level of formal education? (Options: Diploma, Bachelor's, Master's, Doctorate, or Other)

These questions are designed to explore the various aspects of teaching competency and attitudes towards ILTA in vocational settings, categorized into variables that reflect their roles in influencing teaching effectiveness and adaptability. Each question is made to align with the specific influence it is hypothesized to measure, ensuring a comprehensive assessment through the survey. The validation results of the questionnaire are presented in Table 2.

Table 2: Validation of the questionnaire

Variable Type (No. Of Questions)	Variable Description	Sample Questionnaire Item	Cronbach's α
Independent (5)	Professional Development & Technology Use	"I seek professional development opportunities to improve my ILTA."	0.90
Dependent (5)	**Teaching Effectiveness**	"I am confident in my ability to teach ILTA effectively."	0.89
Mediating (3)	**Engagement Strategies**	"I regularly implement strategies that actively engage students in learning ILTA."	0.88
Control (2)	**Experience and Background**	How many years have you been teaching in VE?	N/A

MACHINE LEARNING MODELS EMPLOYED

SVM

SVM is a powerful class of supervised learning algorithms used for classification and regression tasks. The fundamental goal of SVM is to find the optimal hyperplane that maximizes the margin between different classes in the dataset. SVM operates on the principle of finding a hyperplane that best divides a dataset into classes. The best hyperplane is the one that leaves the maximum margin from the nearest points of all the classes. These nearest points are called support vectors. This method is particularly effective in high-dimensional spaces. The main objective of SVM is to solve the following optimization problem:

Objective Function

$$\min_{\mathbf{w},b} \frac{1}{2} \parallel \mathbf{w} \parallel^2 + C \sum_{i=1}^{n} \xi_i \tag{1}$$

Where,

$\mathbf{w}$ is the weight vector of the hyperplane.

b is the bias term.

ξ_i are the slack variables that measure the degree of misclassification of the data x_i.

C is the regularization parameter that controls the trade-off between achieving a low error on the training data and minimizing the norm of the weights to reduce overfitting.

Constraints:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i, \ \forall i \tag{2}$$

where:

y_i are the labels associated with each data point $\mathbf{x}_i$.

The constraints ensure that each data point must lie on or beyond the margin boundary, adjusted by the slack variable ξ_i.

To handle non-linearly separable data, SVM employs the kernel trick, which allows the algorithm to operate in a high-dimensional, implicit feature space without ever computing the coordinates of the data in that space but instead by simply computing the inner products between the images of all pairs of data in the feature space. This is frequently done using a kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$.

Common Kernels

$$\text{Linear: } K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i \cdot \mathbf{x}_j \tag{3}$$

$$\text{Polynomial: } K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i \cdot \mathbf{x}_j + r)^d \tag{4}$$

$$\text{Radial Basis Function (RBF): } K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2\right) \tag{5}$$

where γ, r, and d are parameters that can be tuned according to the data.

Once the optimal values for $\mathbf{w}$ and b are found, the decision function that predicts the class labels is given by:

$$f(\mathbf{x}) = \text{sgn} \left(\sum_{i=1}^{n} y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b\right) \tag{6}$$

where, α_i are the Lagrange multipliers obtained from solving the dual problem associated with the SVM optimization problem.

This decision function allows SVM to classify new data points based on the learned hyperplane and support vectors. The choice of kernel, the value of C, and the kernel parameters (like γ for the RBF kernel) are critical and usually chosen through a cross-validation process to ensure the model generalizes well on unseen data.

Decision Trees (DT)

DT works by creating a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. A tree can be seen as a piece of flowchart-like structure, where each internal node represents a *"test"* on an attribute, each branch represents an outcome of the test, and each leaf node represents a class label or a continuous value.

The process of building a decision tree involves selecting attributes that return the highest information gain or the most considerable reduction in impurity. Here's how this is typically structured:

Node Impurity: At each node, the algorithm chooses the split that results in the most homogeneous sub-nodes. The measure of homogeneity (or its opposite, impurity) is calculated using functions such as:

Gini Impurity (For Classification)

$$I_G(p) = 1 - \sum_{i=1}^{n} p_i^2 \tag{7}$$

Where p_i is the proportion of the samples that belong to class c at a particular node.

Entropy (for classification)

$$I_H(p) = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{8}$$

where p_i is the proportion of the samples that belong to class c at a particular node.

The decision of how to split at each node is based on the measure of impurity. The tree will choose the split that most effectively decreases the impurity of the nodes.

The algorithm will calculate the impurity decrease for every possible split and choose the one that yields the highest gain in purity.

The gain from a potential split is computed as the difference in impurity before the split minus the weighted average impurity of the two child nodes after the split.

To avoid overfitting, DT may be pruned. Pruning reduces the size of the tree by removing parts of the tree that do not provide power in predicting the target variable. One of the most effective pruning techniques is where a complexity parameter α is used to weigh whether nodes can be removed based on the size of the subtree. The trade-off between the subtree's fit to the data and its complexity is calculated, and if the cost is too high, the subtree is pruned.

DTs are straightforward to understand and interpret, making them very useful in operational settings. However, they are prone to overfitting, especially when dealing with very complex trees. This issue can be mitigated through techniques such as pruning and setting limits on tree depth during training. Investigators are able to understand the model's decision-making process because DT is highly visualizable and can deal with quantitative as well as qualitative information. Their success in many different fields can be attributed, primarily in portion, to their openness.

Neural Networks (NN)

NN is obtained from the biological NN that is found in animal brains. From image and speech recognition to Natural Language Processing (NLP), this framework addresses complex pattern recognition and prediction tasks. The network's neurons, or linked nodes, are divided into layers that collectively collect information correlations. An NN's architecture includes many layers, such as input, hidden, and output. Nodes perform

functions and communicate the results to the following layers. Using a sequence of weighted connections, these layers function to convert inputs into outputs.

After introducing a bias and the weighted total of the inputs, each network neuron performs an activation function to analyze the data. This process can be represented as:

$$z^{[l]} = W^{[l]}a^{[l-1]} + b^{[l]}$$

(9)

$$a^{[l]} = g^{[l]}\left(z^{[l]}\right)$$

(10)

where $z^{[l]}$ is the linear combination at layer l, $W^{[l]}$ and $b^{[l]}$ are the weights and biases, $a^{[l-1]}$ is the activation from the previous layer, and $g^{[l]}$ is the activation function.

The choice of activation function in an NN is crucial as it introduces non-linear properties to the system, allowing it to learn more complex patterns. Common activation functions include:

ReLU (Rectified Linear Unit): Useful for hidden layers because it helps with the vanishing gradient problem and speeds up training.

$$\text{ReLU}(x) = \text{Max}(0, x)$$

(11)

Sigmoid: Often used in the output layer of binary classification models, as it maps input to a probability between 0 and 1.

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

(12)

SoftMax: Typically used in the output layer of multi-class classification problems to derive probabilities for each class.

$$\text{SoftMax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

(13)

Learning in neural networks is achieved through a process called backpropagation, where the network adjusts its weights and biases to minimize the error in its predictions.

Loss Functions: A loss function quantifies the difference between the predicted outputs and the actual target values. Common choices include cross-entropy for classification tasks and mean squared error for regression.

Gradient Descent: This optimization algorithm adjusts the weights and biases in a direction that minimally reduces the loss. It computes gradients of the loss function concerning each parameter.

The updated rules for weights and biases can be expressed as:

$$W^{[l]} = W^{[l]} - \alpha \frac{\partial \text{Loss}}{\partial W^{[l]}}$$

(14)

$$b^{[l]} = b^{[l]} - \alpha \frac{\partial \text{Loss}}{\partial b^{[l]}}$$

(15)

where α is the learning rate, a hyperparameter that controls the step size during the learning process.

NN is extremely flexible and powerful, capable of modeling highly intricate relationships due to their deep and complex architectures. Yet they are frequently referred to as "black boxes" because of the challenges involved in figuring out how the model functions and the massive amount of data and computing resources required for effective instruction.

DATA PREPROCESSING

Data cleaning is one step in data preprocessing; it involves eliminating any errors, missing numbers, or anomalies from the dataset. Statistical techniques like mean imputation or more complex ones like regression matching might be employed to fill in missing values in data from this investigation, whether it's responses to questionnaires or performance test results. Mode approximation is frequently employed for qualitative information like the teachers' educational histories.

When data covers units and scales, feature scaling is a vital procedure after cleaning. This is especially true for NN, which are susceptible to input scales because they have been optimized using gradient descent. Methods such as standardization, which adapts data such that it has a mean of zero and a standard deviation of one, and normalization, which changes data such that it matches within a scaled range like 0 to 1, are used. For the learning procedure to be statistically stable and to avoid any one feature from controlling the model because of its size, this standardization is important.

To make statistical types accessible by the model, qualitative data is further processed using feature encoding. Because of the fundamental classification of data like higher education or types of VE, this is of the highest priority. As a prime instance of a usual procedure, one-hot encoding involves changing particular variables into a syntax that ML algorithms can use to enhance prediction.

The use of feature engineering to develop new features is an additional vital component of preliminary processing. More knowledge or predictive ability could be obtained from existing data by extracting new variables. Examine the idea of developing a "experience index" that takes into account both the number of years worked teaching and more limited qualifications in order to provide a more unambiguous indication of an instructor's ability to address challenging subjects like ILTA.

Data is separated into training and test sets after preliminary processing is complete. It is typical to train the model using an enormous amount of the data, around 70%, and then test how it uses the remaining 30%. In order to test the model on fresh data and make sure the trends it learned are valid outside of the known data, this split is necessary. Table 3 illustrates the preliminary processing steps' impact on each key feature, both before and after:

Table 3: Preprocessing process

Data Attribute	Sample Before Processing	Processing Technique	Sample After Processing
Years of Teaching	Null	Mean Imputation	10 (assumed mean)
Educational Background	Master's degree	One-hot Encoding	[0, 0, 1, 0, 0] (assuming order: Diploma, Bachelor's, Master's, Doctorate, Other)
Teaching Effectiveness Score	85 (out of 100)	Normalization	0.85 (scaled between 0-1)
Type of Training	Culinary Arts	Label Encoding	3 (assuming arbitrary mapping)
Age	46	Standardization	0.5 (example standardized value)

EXPERIMENTAL SETUP

Software and hardware tuned to fulfill the requirements of the ML models drive the setup used for experiments. It includes a 64 GB DDR4 RAM, an NVIDIA Tesla P100 GPU with 16 GB VRAM, an Intel Xeon CPU E5-2630 v4 @ 2.20 GHz, and a 1 TB SSD for memory. Python 3.8 is used by the system, along with libraries like Scikit-learn for ML, TensorFlow for deep learning (DL), and the Pandas program to perform data processing. This arrangement ensures the effectiveness of teaching and assessing the models. The hyperparameters which are defined in Table 4 are used to train the models.

Table 4: Hyperparameter Table

Model	Hyperparameter	Description	Value
	C	Regularization parameter	1
SVM	Kernel	Specifies the kernel type to be used in the model	RBF
	Gamma	Kernel coefficient for RBF	0.1

DT	Max Depth	The maximum depth of the tree	10
	Min Samples Split	The minimum number of samples required to split a node	20
	Criterion	The function to measure the quality of a split	Gini
NN	Number of Layers	Number of hidden layers	2
	Neurons per Layer	Number of neurons in each hidden layer	50
	Activation Function	Activation function for hidden layers	ReLU
	Learning Rate	Step size at each iteration of the learning process	0.01
	Loss Function	The objective function to be minimized	Cross-Entropy

Parameters like Accuracy, Precision, Recall, F1-score, ROC, and confusion matrix are employed to evaluate the performance of ML models.
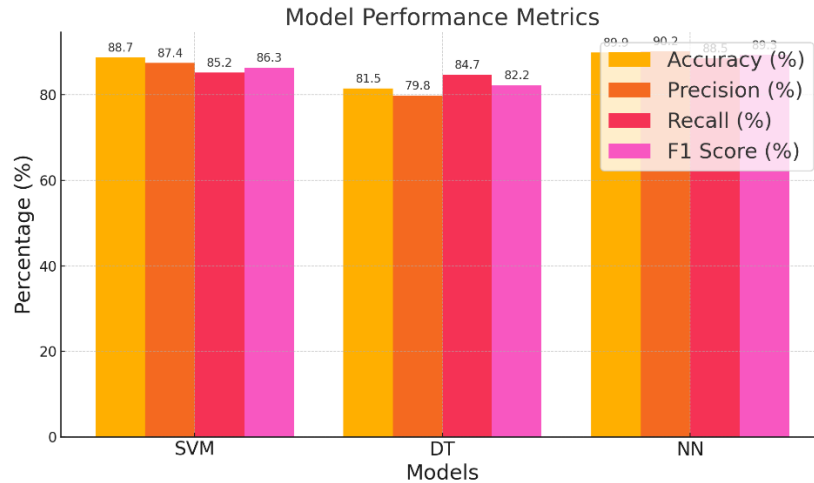

Figure 1: Performance analysis for different metrics

Figure 1 presents the outcomes for each metric for the three ML models.

SVM: There is a significant amount of positive predictive value and comparatively few False Positives (FP) when using SVM due to its high precision and accuracy. However, it has fewer recalls, indicating that it could miss some positive cases (also called false negatives, or FN). The balance between recall and precision manifests in the F1 score which is balanced.

DT: The DT's inadequate results in terms of accuracy and precision may be related to its susceptibility to overfitting, which becomes more evident when handling large or noisy datasets. It collects more positive cases but is at risk of more false positives (FP) owing to its relatively good recall.

NN: NN exhibits the highest accuracy and precision, indicative of robust performance across most cases with minimal FP and FN. The model's prudent thresholding in class prediction, resulting in certainty above protection, could explain why the recall is slightly lower than the precision.
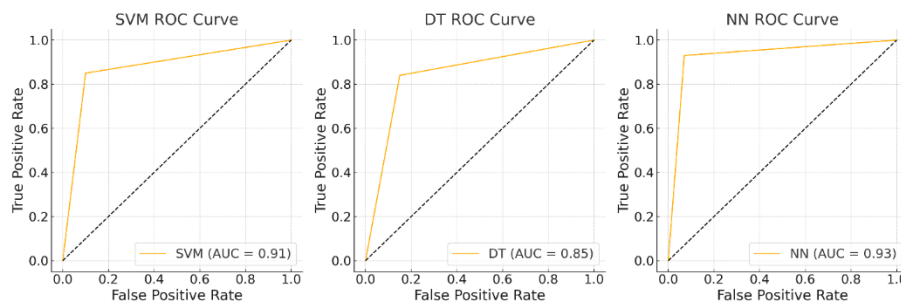

Figure 2: RoC for three models

Figure 2 demonstrates how well each of the three ML models scored on the ROC test.

SVM: With a reduced likelihood of FP and FN, an AUC of 0.91 indicates a high capacity to differentiate between the positive and negative classes. This result shows that the SVM model is superior to chance when it comes to ranking forecasts.

DT: While positive, DT's AUC of 0.85 suggests the model may have some drawbacks when it comes to class distinction. The model's overfitting risk or its simple decision limits could be responsible if they do not recognize the complex nature of data distributions that are more complicated or have overlapped.

NN: With an AUC of 0.93, the NN demonstrates excellent classification performance, likely due to its DL capabilities that excel in capturing intricate patterns in the data. This high score reflects its strength in handling varied data inputs and its robustness in generalization.
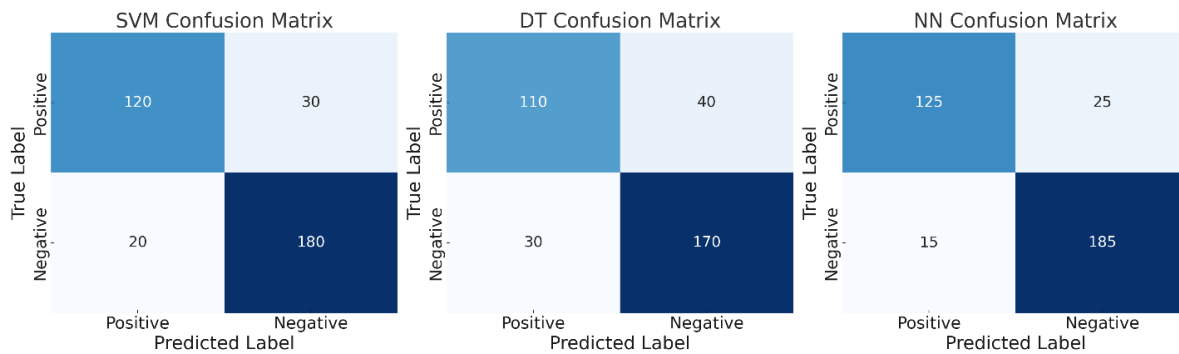

Figure 3: Confusion matrix for three models

Figure 3 presents the confusion matrix for the three models.

SVM Confusion Matrix: This model demonstrates a strong performance with 120 True Positives and 180 True Negatives (TN), indicating a high ability to identify both positive and negative cases correctly. The model recorded 30 False Positives and 20 FN, suggesting that while it is reliable, there are some instances where it either over-predicts the positive class or fails to detect it. The relatively low number of FN and FP shows a balanced trade-off between sensitivity (Recall) and specificity.

DT Confusion Matrix: The DT shows slightly lower efficacy with 110 TP and 170 TN. There are 40 FN and 30 FP, indicating that this approach may have had more difficulty determining both classes than SVM. The higher rate of FP and FN shows that the DT model may be generating additional errors in class prediction due to overfitting the training data or weak generalization.

NN Confusion Matrix: Among the three models, the NN scores most successfully with 125 TP and 185 TN. Its 25 FP and 15 FN are the most affordable of any of the options. Both the accuracy with which positive cases are classified and the reliability with which negative cases are ignored are strongly recommended. The model's ability to capture a high percentage of positive cases with no loss of precision for recall is shown by a small amount of FP.

CONCLUSION

These findings of the present investigation demonstrate that ML algorithms are excellent at measuring vocational teachers' Information Literacy Teaching Ability (ILTA) wholly and independently. While analyzing different aspects of a teacher's performance, each of the three models (SVM, DT, and NN) displayed a unique set of advantages. The superior precision and generalizability shown by NN most notably testify to their possibility as a successful tool for academic assessments. This study not only provides weight to the possibility of using ML in educational assessments but also provides a path for further investigations into particular educational performance factors. In addition, by examining the data, we may enhance the quality of VE through focused professional development initiatives and curriculum changes. Using data-driven techniques, VE institutions can adapt their curriculum to meet the evolving requirements of current employees while

maintaining the rapid development of technological advancement. More educated and successful methods of instruction that can adapt to the educational hurdles of the 21st century will be feasible because of this study, which provides an initial step toward modernizing academic evaluation practices in vocational environments.

REFERENCES

Gyimah, N. (2020). Assessment of Technical and Vocational Education and Training (TVET) on the development of the World's Economy: Perspective of Africa, Asia and Europe. Asia and Europe (February 19, 2020).

Ling, Y., Chung, S. J., & Wang, L. (2023). Research on the reform of management system of higher vocational education in China based on personality standard. Current Psychology, 42(2), 1225-1237.

Ha, W., Yang, P., Choi, Y., Ra, S., Hayashi, R., & McCutcheon, C. (2023). Crossing the rivers by touching the stones: Alternative approaches in technical and vocational education and training from the People's Republic of China and the Republic of Korea. ECNU Review of Education, 6(4), 654-676.

Darder, A., Hernandez, K., Lam, K. D., & Baltodano, M. (2023). Critical pedagogy: An introduction. In The critical pedagogy reader (pp. 1-30). Routledge.

Phippen, A., Bond, E., & Buck, E. (2021). Practical strategies for information literacy education: combatting 'fake news' and empowering critical thinking. In Future Directions in Digital Information (pp. 39-53). Chandos Publishing.

Hammons, J. (2020). Teaching the teachers to teach information literacy: A literature review. The Journal of Academic Librarianship, 46(5), 102196.

Carpenter, S. K., Witherby, A. E., & Tauber, S. K. (2020). On students'(mis) judgments of learning and teaching effectiveness: Where we stand and how to move forward.

Hammons, J. (2020). Teaching the teachers to teach information literacy: A literature review. The Journal of Academic Librarianship, 46(5), 102196.

Norrie, C. S., Deckers, S. R., Radstaake, M., & van Balkom, H. (2024). A Narrative Review of the Sociotechnical Landscape and Potential of Computer-Assisted Dynamic Assessment for Children with Communication Support Needs. Multimodal Technologies and Interaction, 8(5), 38.

B. Lin, "A Study on College English Teachers' Informatization Teaching Ability in Private Colleges Based on Information Technology in Blended Teaching Mode," 2021 2nd International Conference on Information Science and Education (ICISE-IE), Chongqing, China, 2021, pp. 259-263, doi: 10.1109/ICISE-IE53922.2021.00067.

L. Zhang, J. Yang, Y. Li, T. Wu and F. Wu, "An investigation on the information literacy of primary and secondary school teachers in one Province of Western China," 2021 Tenth International Conference of Educational Innovation through Technology (EITT), Chongqing, China, 2021, pp. 330-333, doi: 10.1109/EITT53287.2021.00071.

H. Lei, "Informational Teaching Concept in the Reform of College English Education," 2020 International Conference on Computers, Information Processing and Advanced Education (CIPAE), Ottawa, ON, Canada, 2020, pp. 145-147, doi: 10.1109/CIPAE51077.2020.00046.

M. Chen, C. Zhou and Y. Wu, "Research on the Model and Development Status of Information Literacy Self-improvement Ability of Primary and Secondary School Teachers," 2020 Ninth International Conference of Educational Innovation through Technology (EITT), Porto, Portugal, 2020, pp. 87-91, doi: 10.1109/EITT50754.2020.00021.

L. Yanfen and H. Guanwen, "Research on the Current Situation and Countermeasures of Teachers' Information Literacy Based on the Information Technology—Take Guangzhou's Higher Vocational Tourism Teachers as an Example," 2021 2nd International Conference on Big Data and Informatization Education (ICBDIE), Hangzhou, China, 2021, pp. 696-700, doi: 10.1109/ICBDIE52740.2021.00164.

J. Ang and H. Zhang, "Investigation and Analysis of Online Teaching in Higher Vocational Colleges during the COVID-19 Epidemic," 2021 9th International Conference on Information and Education Technology (ICIET), Okayama, Japan, 2021, pp. 252-256, doi: 10.1109/ICIET51873.2021.9419577.