Enhancing Competency Level Prediction Using Machine Learning: A Data-Driven Approach Based on Psychological Assessment Data

Sinung Suakanto[1*], Joko Siswanto[2], Jan M. Pawlowski[3], Muharman Lubis[4], Syfa Nur Lathifah[5], and Litasari Widyastuti Suwarsono[6]
[1,4,5]Information System, School of Industrial Engineering, Telkom University (Main Campus)
Jawa Barat, Indonesia 40257
[2]Industrial Engineering, Faculty of Industrial Technology, Bandung Institute of Technology
Jawa Barat, Indonesia 40132
[3]Business Information Systems, Institute of Computer Science,
Ruhr West University of Applied Sciences
Bottrop, Germany 46236
[6]Digital Psychology, School of Communication and Social Sciences, Telkom University
(Main Campus)
Jawa Barat, Indonesia 40257
Email: [1]sinung@telkomuniversity.ac.id, [2]j.siswanto@itb.ac.id, [3]Jan.Pawlowski@hs-ruhrwest.de, [4]muharmanlubis@telkomuniversity.ac.id, [5]syfanr@telkomuniversity.ac.id, [6]litasari@telkomuniversity.ac.id,

***Abstract*—Competency level prediction plays a crucial role in competency-based human resource management such as talent management. Talent management is achieved by identifying individuals' knowledge, skills, and attitudes through psychological assessment. Recognizing employees as a strategic asset by accurately predicting competencies supports targeted development, boosting individual and organizational performance. Current practices related to competency assessment require expert judgment from psychologists or assessors, which can be time-consuming. The research proposes a machine learning–based approach to predict competency levels using psychological assessment scores as input, designed to operate within digital, network-enabled interview platforms. Several machine learning methods, including Random Forests, k-Nearest Neighbors (KNN), and Support Vector Machines (SVMs), are applied to historical assessment datasets to identify patterns and relationships between psychological assessment scores and competency levels.The dataset comprises 1,220 records from a psychological assessment. The experimental results indicate that the Random Forest model achieves the highest accuracy of 81%, outperforming other models in competency level prediction. The key novelty lies in its data-driven methodology, which enhances the objectivity and efficiency of competency evaluation while reducing reliance on expert interpretation. By enabling automated competency prediction in network-enabled interview environments, the proposed approach supports more efficient talent decision-making, workforce development, and recruitment processes. The findings demonstrate that machine learning can accurately predict competency levels from a clean dataset of psychological assessment scores, achieving accuracy above 80%. Future research may enhance model robustness by incorporating additional assessment center criteria and real-world performance metrics.**

***Index Terms*—Competency Level Prediction, Machine Learning, Psychological Assessment, Human Capital as Strategic Asset, Network-Enabled Interview Platform**

Received: May 24, 2025; received in revised form: Sep. 08, 2025; accepted: Sep. 08, 2025; available online: April 02, 2026.
*Corresponding Author

I. INTRODUCTION

COMPETENCY is a critical component in talent management, encompassing the integrated knowledge, skills, and behaviors required to perform a job effectively [1–3]. As organizations strive for excellence, the need for accurate assessment in developing these competencies becomes increasingly urgent [4]. From a strategic perspective, employees are among the most valuable organizational assets, and enhancing their competencies directly contributes to sustaining competitive advantage and long-term business value. Competency assessments are crucial for identifying employees' current skill levels and improving their capabilities to support organizational goals and future demands [5].

Complementing competencies, potential refers to an individual's inherent ability to acquire new skills over time, often measured through cognitive abilities, personality traits, and learning agility [2, 3]. Potential is described as "hidden talents and fundamental skills of an individual that have not yet been fully realized" [6]. When nurtured through experience and training, this potential contributes significantly to long-term organizational success [7].

Assessing both competency and potential requires certified professionals. Psychologists evaluate latent attributes such as cognitive ability and emotional intelligence [8], while certified assessors evaluate competency levels by assessing how well individuals meet specific job requirements [9, 10]. This two-tiered evaluation ensures accurate and holistic talent assessments, enabling organizations to make informed talent management decisions [11].

Despite its importance, competency assessment poses challenges, especially for long-tenured employees. Evaluations are often limited to initial recruitment stages using psychological assessment tools [12–14], whereas ongoing competency reassessment is often resource-intensive, time-consuming, and costly. Moreover, existing methods, such as self-assessments or managerial reviews, are prone to bias and lack scalability [15].

Psychological assessment data also remain widely used to measure latent attributes, such as cognitive abilities, personality traits, and emotional intelligence, providing standardized and quantifiable data [16]. However, converting these results into accurate competency assessments is complicated by evaluator bias, subjectivity, and the complex nature of competency itself. Traditional techniques, including self-assessments and managerial reviews, are subject to biases, such as self-serving bias, which compromise assessment accuracy [13]. Although expert-led behavioral simulations and interviews improve accuracy, these processes are laborious and impractical for large workforces [12–14]. Therefore, there is a critical need for innovative, data-driven tools to enhance competency evaluation.

In response to these challenges, the research proposes the development of an AI-based system as a productivity tool to streamline the assessment process and reduce the workload on human resources professionals by automating various aspects of competency evaluation. This efficiency is crucial in today's fast-paced work environment, where timely and precise assessments can significantly impact organizational performance. The proposed system leverages machine learning to estimate employee competencies using psychological assessment data as proxies for potential. Machine learning has shown a good result in predicting employee performance [3, 16–19] and competencies across various domains such as nursing [20], construction [21], education [22], and marketing [23]. The previous studies have used diverse inputs including psychological assessment scores, performance history, demographic data, and other relevant personal and work-related factors. In human resource contexts, machine learning has also been used to forecast behaviors such as turnover [24], attrition [25], and job involvement [26], as well as to assess person-job fit [27–29]. Psychological assessment data serve as valuable predictors in machine learning models [16, 30]. These psychological assessment data may include cognitive ability, job knowledge, such as emotional intelligence, conscientiousness, and extraversion, assessment center scores, integrity, biodata, interview ratings, Situational Judgment Test (SJT) scores, Predictive Occupational fit (POfit), and vocational interests [16]. Table I shows the summary of previous studies.

Prior studies have explored both statistical and algorithmic approaches to predict competency [1–3, 34–38]. Their findings indicate that machine learning outperforms traditional methods in handling complex, high-dimensional datasets. Thus, this research builds on and reinforces the growing body of evidence supporting the use of advanced machine learning techniques over conventional methods for modeling the intricate relationships between psychological assessment variables and competency outcomes. The research aims to build and validate a machine learning model that predicts competency level using psychometric assessment data, thereby bridging the gap between latent potential and observable competency. By training the model on historical datasets that combine psychological assessment inputs with competency evaluation outcomes, this approach aspires to reduce subjectivity, enhance assessment efficiency, and provide a scalable, data-driven solution for talent management.

The contributions of the research are significant for recruitment, professional development, and organizational planning. By enabling more accurate competency predictions, the proposed tool supports personalized learning strategies, informed decision-making, and optimized workforce deployment, ultimately improving organizational performance and competitiveness. The research focuses three main research questions:

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, “Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data”, CommIT Journal 20(1), 139–154, 2026.

TABLE I
RESEARCH STATE OF THE ART

Author	Predicted Variable	Input Variables	Model	Accuracy
Job/Employee Performance				
[3]	Job performance (Excellent / Other)	Individual competencies (knowledge, skills, roles, values, attitudes, and motivation)	GA-BP Neural Network	91.16%
[17]	Employee performance	Employee performance data (attendance, evaluations, training, demographics, and work history)	Linear Regression, DT, RF, SVM	–
[18]	Employee performance (low/medium/high)	Employee attributes grouped into general, physical, behavioral/social, and economic factors	LR, GNB, DT, KNN, SVM	RF: 98.2%
[19]	Employee performance (high-/moderate/low)	Historical employee data (demographics, education, role, tenure, performance, attendance)	Various machine learning models	–
[16]	Job performance	Psychometrically validated test scores	OLS vs Random Forest	Modern machine learning better
Competency				
[20]	Community nursing workability	Nursing staff abilities and personal attributes	R-GCN-GRU	98.4%
[30]	Employee competence (4 classes)	Employee assessment and learning data (3,634 records)	NB, KNN	KNN: 99.33–99.45%
[31]	Behavioral competencies	Public Instagram profile features	GLM, DL, RF	69–70%
[21]	Safety competency	Organizational network, behavioral, social capital, and experience indicators	GBDT	–
[22]	Open education competency	Student perceptions of open education competencies (eOpen)	DT, RF	RF: 93.94%
[23]	Selling competency (yes/no)	Bank telemarketing campaign attributes	J48	91.2%
Employee Behaviour (Turnover/Attrition/Discipline)				
[24]	Employee turnover	Employee demographic and job-related attributes	RF, Lasso	AUC > 0.8
[32]	Turnover intention	Graduate employment survey data (general, physical, behavioral, and economic factors)	LR, KNN, XGB	XGB: 78.5%
[25]	Employee attrition	IBM HR attrition dataset	LR, CT, RF, NB, NN, Ensemble	LR: 88%
[26]	Disciplinary violation	Employee demographics, job details, and disciplinary records	RF, DT, NB, Ensemble	RF: 87.30%
[33]	Job involvement (4 levels)	Employee demographics, job characteristics, performance, satisfaction, and compensation	GLM	67.69%
Person–Job/Skill Fit				
[27]	Person–job fit	Job application data (resumes and job requirements)	CNN / PJFNN	85% (AUC)
[28]	Job title classification	Employees’ demographic and employment records (5,000 employees)	KNN, RF, GB, SVM	KNN: 96%
[29]	Skill shortage classification	Labor demand and workforce indicators	XGBoost	83%

Note: Genetic Algorithm–Backpropagation (GA-BP), Decision Trees (DT), Random Forest (RF), Support Vector Machines (SVM), Linear Regression (LR), Gaussian Naive Bayes (GNB), K-Nearest Neighbors (K-NN), Relational Graph Convolutional Network - Gated Recurrent Unit (R-GCN-GRU), Ordinary Least Squares (OLS), Gradient Boosting Decision Tree (GBDT), Naive Bayes (NB), Generalized Linear Model (GLM), Deep Learning (DL), Area Under the Curve (AUC), Extreme Gradient Boosting (XGB), Classification Tree (CT), Neural Network (NN), Convolutional Neural Network (CNN), Person-Job Fit Neural Network (PJFNN), and Gradient Boosting (GB).

(1) how machine learning models are designed to effectively predict employee competencies using psychological assessment data as proxies for potential, (2) which psychological assessment variables are the most significant predictors of competency outcomes in the proposed model, and (3) how the predictive performance of the machine learning model compares to traditional competency assessment methods in terms of accuracy, scalability, and bias reduction. In practical organizational settings, such machine-learning–based competency-prediction models are increasingly deployed on digital, network-enabled interview platforms to support scalable, remote assessment processes. Therefore, the research also presents a system-level perspective to illustrate how the proposed model can be integrated into an automated Interview-Bot environment, while maintaining the primary focus on predictive model development and evaluation.

II. Research Method

To contextualize the proposed machine learning–based competency prediction approach, the research presents a conceptual framework that captures how the model operates in a digital, network-enabled interview environment. In contemporary human resource assessment practices, psychological evaluations

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, "Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data", CommIT Journal 20(1), 139–154, 2026.

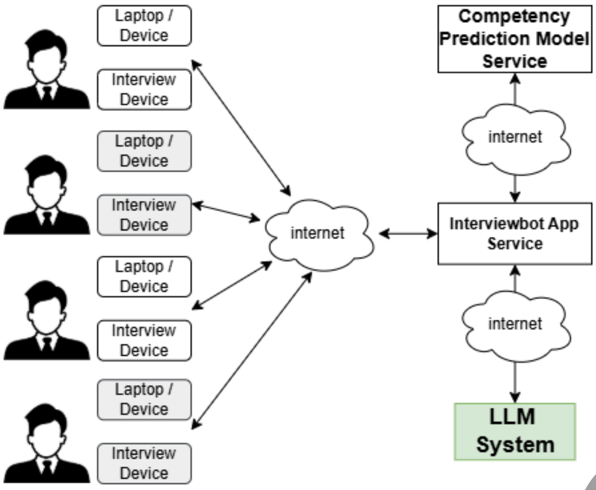

Fig. 1. Conceptual architecture of the network-enabled InterviewBot system. Note: Large Language Model (LLM).

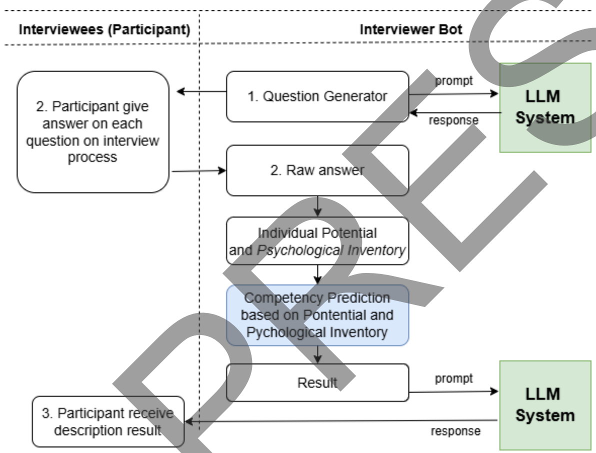

Fig. 2. Interview process flow for machine learning–based competency prediction. Note: Large Language Model (LLM).

are increasingly conducted through distributed interview platforms that support multiple concurrent users accessing the system using laptops or dedicated interview devices. Figure 1 illustrates the high-level conceptual framework, highlighting the interactions among interview participants, the InterviewBot application, network infrastructure, and backend intelligent services that enable scalable, remote interview processes.

The competency prediction model functions as the core analytical component of the InterviewBot. As illustrated in Fig. 2, the InterviewBot manages the interview flow by generating structured interview questions, collecting participants' responses, and transforming raw responses into psychological inventory indicators. These indicators are processed by the machine learning–based competency prediction model to estimate competency levels. In addition, a Large Language Model (LLM) is utilized to support dynamic question generation and to produce interpretable competency descriptions, enabling the InterviewBot to deliver meaningful feedback to participants.

Following the system-level and interaction-level descriptions, the research focuses on developing a machine learning model for competency prediction. To build and evaluate the proposed model, a conceptual prediction framework is illustrated in Fig. 3. This framework describes how psychological assessment data are processed as input features and mapped to multiple competency level outputs using a multi-class classification approach. The conceptual prediction model serves as the methodological foundation for implementing, training, and evaluating the machine learning algorithms discussed in the subsequent sections.

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, "Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data", CommIT Journal 20(1), 139–154, 2026.

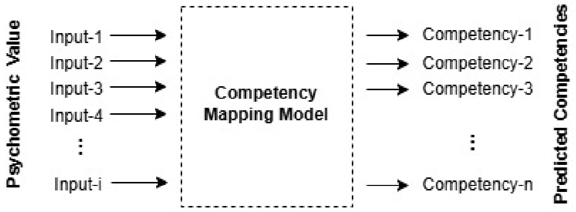

Fig. 3. Conceptual framework for competency prediction. Note: Machine Learning (ML).

The dataset includes respondents from various genders and functional roles within the organization, with no restrictions on age, position, or tenure. All available records from the psychological assessments are considered for inclusion. Records are excluded if they are incomplete (missing values in any input variable), contain duplicate entries, or show extreme outliers beyond three standard deviations from the mean in key variables. After applying these criteria, only clean and complete records are retained for model development.

The research method framework shown in Fig. 4 illustrates the processes of data preparation, model training, and testing. Pre-processing steps include (1) removing duplicate entries to avoid redundancy, (2) handling of missing values by discarding incomplete records, (3) detecting outliers and removing data using z-score thresholds, (4) normalizing continuous variables to ensure scale uniformity, and (5) encoding of categorical variables into numerical format for model compatibility. This process ensures that the dataset is consistent, clean, and suitable for machine learning algorithms.

During the training phase, the dataset undergoes several steps, including pre-processing, model training using various machine learning methods (model training), and hyperparameter optimization (tuning) to yield a trained model. Subsequently, in the testing phase, this trained model is evaluated using unseen new data (new or test data) through a model testing process to generate predicted data. The evaluation measures outcomes by comparing predicted data with the actual ground truth of the test data. The evaluation of the model's performance uses key metrics, such as accuracy (overall correctness), precision (the ratio of true positives to all predicted positives), recall (the ratio of true positives to all actual positives), and the F1-Score (the harmonic mean of precision and recall).

The input variables include psychological assessment scores across various cognitive abilities, personality traits, and indicators of emotional intelligence. A detailed list of these input variables (39 input variables) is provided in Table II. The psychological assessment data sources (input variables) include the Holland Type

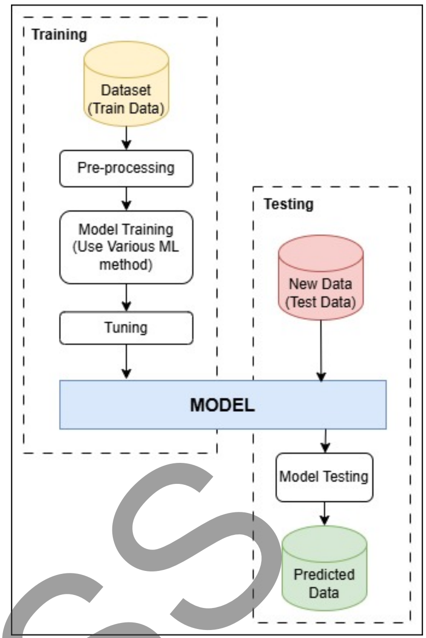

Fig. 4. Research method. Note: Machine Learning (ML).

I & II, the Differential Aptitude test, the 3 Competence - Commitment - Character (3C) assessment, and the Personality and Preference Inventory (PAPI) Kostick.

Meanwhile, the output variables are competency levels, which reflect the organization's competency model and are categorized according to established competency frameworks related to the Spencer competency model [39]. Although the Spencer competency framework is introduced decades ago, it remains widely used today as a key reference for defining and assessing competencies [40–43]. The list of output variables is shown in Table III.

In the Spencer competency model [39], each competency is assessed on a five-level scale, ranging from 0 to 4. Each level is associated with specific behavioral indicators that describe observable actions corresponding to the degree of competency mastery. The levels are as follows. Level 0 does not demonstrate this competency or only shows it in a very limited capacity. Level 1 demonstrates competency in simple or basic situations. Level 2, consistently, applies the competency in standard work situations. Level 3 demonstrates the competency in more complex situations and is capable of guiding others. Level 4 shows the competency at a strategic level, influencing organizational policies and contributing to long-term development. Following this, model development is carried out by testing multiple machine learning algorithms using seven methods, as shown in Table IV.

To optimize model performance, hyperparameter

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, "Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data", CommIT Journal 20(1), 139–154, 2026.

TABLE II
INPUT VARIABLES AND PSYCHOMETRIC ASSESSMENT TOOLS.

No.	Code	Variable	Psychometric Assessment Tools Source
1	H1-R	R – Realistic	Holland Type I – Personality Assessments
2	H1-I	I – Investigative	Holland Type I – Personality Assessments
3	H1-A	A – Artistic	Holland Type I – Personality Assessments
4	H1-S	S – Social	Holland Type I – Personality Assessments
5	H1-E	E – Enterprising	Holland Type I – Personality Assessments
6	H1-C	C – Conventional	Holland Type I – Personality Assessments
7	H2-R	R – Realistic	Holland Type II – Personality Assessments
8	H2-I	I – Investigative	Holland Type II – Personality Assessments
9	H2-A	A – Artistic	Holland Type II – Personality Assessments
10	H2-S	S – Social	Holland Type II – Personality Assessments
11	H2-E	E – Enterprising	Holland Type II – Personality Assessments
12	H2-C	C – Conventional	Holland Type II – Personality Assessments
13	DAT2S	Differential Aptitude Test (DAT) Type II	Differential Aptitude Test
14	DAT3S	Differential Aptitude Test (DAT) Type III	Differential Aptitude Test
15	DAT4S	Differential Aptitude Test (DAT) Type IV	Differential Aptitude Test
16	DAT5S	Differential Aptitude Test (DAT) Type V	Differential Aptitude Test
17	DAT6S	Differential Aptitude Test (DAT) Type VI	Differential Aptitude Test
18	DAT7S	Differential Aptitude Test (DAT) Type VII	Differential Aptitude Test
19	DAT8S	Differential Aptitude Test (DAT) Type VIII	Differential Aptitude Test
20	DAT9S	Differential Aptitude Test (DAT) Type IX	Differential Aptitude Test
21	DAT10S	Differential Aptitude Test (DAT) Type X	Differential Aptitude Test
22	Cor	Core	3C Assessment
23	Cal	Calculation	3C Assessment
24	Cog	Cognition	3C Assessment
25	ACH	Achievement	PAPI Kostick (Personality and Preference Inventory)
26	DEF	Defensiveness	PAPI Kostick
27	ORD	Orderliness	PAPI Kostick
28	EXH	Exhibitionism	PAPI Kostick
29	AUT	Autonomy	PAPI Kostick
30	AFF	Affiliation	PAPI Kostick
31	INT	Introspection	PAPI Kostick
32	SUC	Succorance	PAPI Kostick
33	DOM	Dominance	PAPI Kostick
34	ABA	Abasement	PAPI Kostick
35	NUR	Nurturance	PAPI Kostick
36	CHG	Change	PAPI Kostick
37	END	Endurance	PAPI Kostick
38	HET	Heteronomy	PAPI Kostick
39	AGG	Aggression	PAPI Kostick

TABLE III
OUTPUT VARIABLES.

Code	Competency	Alias Name (Spencer Model)
ACH	Result Orientation	Achievement Orientation
ING	Integrity	Integrity
COL	Collaboration	Teamwork & Cooperation
ADP	Adaptability	Flexibility
CLR	Continuous Learning	Self-Development
VD	Valuing Diversity	Interpersonal Understanding
CO	Attention to Quality & Clarity of Tasks	Concern for Order, Quality & Accuracy
INF	Information Seeking	Information Seeking
IU	Empathy	Interpersonal Understanding
SO	Service Orientation	Customer Service Orientation
IMP	Influencing Others	Impact & Influence
OA	Organizational Awareness	Organizational Awareness
DIR	Directing Ability	Directiveness
LD	Leadership	Leadership
AT	Analytical Thinking	Analytical Thinking
CT	Conceptual Thinking	Conceptual Thinking
SCT	Self-Control	Self-Control
OC	Organizational Commitment	Commitment to Organization

TABLE IV
PROPOSED MACHINE LEARNING METHODS.

No.	Method	Description
1	Gradient Boosting	Combines sequential trees to reduce prediction errors [44].
2	Random Forest	Aggregates multiple randomly generated decision trees [45].
3	Support Vector Machine	Maximizes the margin between class boundaries [46].
4	Neural Network	Learns patterns through layered node activations [47].
5	K-Nearest Neighbors	Classifies instances based on nearest neighbors [48].
6	Decision Tree	Splits data using feature-based decision rules [49].
7	AdaBoost	Combines weak learners by reweighting misclassified samples [50].

tuning is conducted for each algorithm to identify the most suitable configuration, thereby improving prediction accuracy and model stability. Model evaluation is carried out using several standard performance metrics, including the confusion matrix, accuracy, precision, recall, and F1-Score, to comprehensively assess the classification performance. Furthermore, the predictive performance of different machine learning models is systematically compared to determine the most effec-

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, “Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data”, CommIT Journal 20(1), 139–154, 2026.

tive approach for competency level prediction based on psychological assessment data. This comparative analysis provides a deeper understanding of how different algorithms capture patterns within the dataset and handle complex relationships among psychological variables. In addition, the evaluation process ensures the robustness and reliability of the proposed predictive framework. The results also offer insights into the relative strengths and limitations of each model in supporting data-driven competency assessment.

TABLE V
FEATURE IMPORTANCE FOR COMPETENCY CONCEPTUAL THINKING (CT).

Rank	Feature(s)	Importance	Interpretation
1	DAT7S	0.113	The 7th Differential Aptitude Test (DAT) subtest is the most influential variable in predicting CT. It is likely related to abstract reasoning or a similar cognitive domain.
2–5	DAT5S, DAT4S, DAT6S, DAT8S	0.07–0.09	Other DAT subtests also play a significant role, which makes sense since CT is strongly related to abstract thinking, logic, and problem-solving abilities.
6–7	DAT3S, DAT10S	0.05–0.06	It is still within the range of DAT cognitive subtests, contributing moderately to CT prediction.
8	H1-I (Investigative)	0.03	Holland type “Investigative” indicates that individuals with high intellectual curiosity tend to exhibit stronger Conceptual Thinking.
13	Change (CHG)	0.016	From the PAPI Kostick test, people who enjoy change (adaptive and innovative) are likely to support stronger CT.
14	Orderliness (ORD)	0.016	Being orderly and systematic may contribute positively to conceptual thinking abilities.

III. Results and Discussion

This section presents the study’s findings, including data exploration results, model performance comparisons, and an in-depth discussion of the selected machine learning model. The initial stage of data exploration provides insights into the dataset’s structure, distributions, and relationships between psychological assessment scores and competency outcomes. Subsequently, multiple machine learning models are trained and evaluated to compare their predictive performance using several evaluation metrics. Finally, the best-performing model is analyzed in greater detail, including the interpretation of feature importance to understand the key psychological factors influencing competency level prediction.

A. Dataset

The dataset used comprises 1,220 records from a psychological assessment. The data are collected from a large organization with a substantial workforce of approximately 230 employees. For labeling, competency scores are obtained from assessments conducted by certified psychologists and professional assessors. To minimize potential bias and ensure consistency in the labeling process, the dataset is limited to a single organization with a diverse internal employee population. This dataset includes results from psychological assessments that generate psychological assessment scores, as well as competency scores derived from two sources: evaluations conducted by a team of psychologists and specialized competency assessments.

Before developing the model, an initial preprocessing stage is performed to clean the dataset. This step involves handling missing values and outliers, which are primarily caused by incomplete assessments or system errors. Due to these inconsistencies, a significant portion of the dataset has to be excluded. After data cleaning, only 221 of the initial 1,220 entries are retained for further analysis and model training. This reduction is primarily due to missing values (80%) and the removal of extreme outliers, with a small number of duplicate records also excluded. No specific inclusion or exclusion criteria related to demographic characteristics are applied. The dataset comprises both male and female respondents from various functional roles and job positions. Given this diversity, no formal categorization of respondent characteristics is conducted.

The remaining sample size of 221 is considered sufficient to support model development with 39 predictor variables. The distribution of output variables for each class label (0–4) is presented in Table V. These figures illustrate the frequency of instances in each competency class. Certain competencies, such as service orientation (SO), exhibit skewed distributions toward specific levels, which may reduce prediction accuracy.

B. Data Exploration

In this section, a correlation analysis is conducted to examine the relationships among the input variables and the output variable. The first step is to analyze the correlations among input variables, such as cognitive ability, personality traits, and emotional intelligence, to identify potential interdependencies. This correlation matrix is shown in Fig. 5. There are strong positive correlations (indicated by deep red) along the diagonal, as expected since each variable is perfectly correlated with itself. There are also notable positive correlations (lighter shades of red) between certain groups of vari-

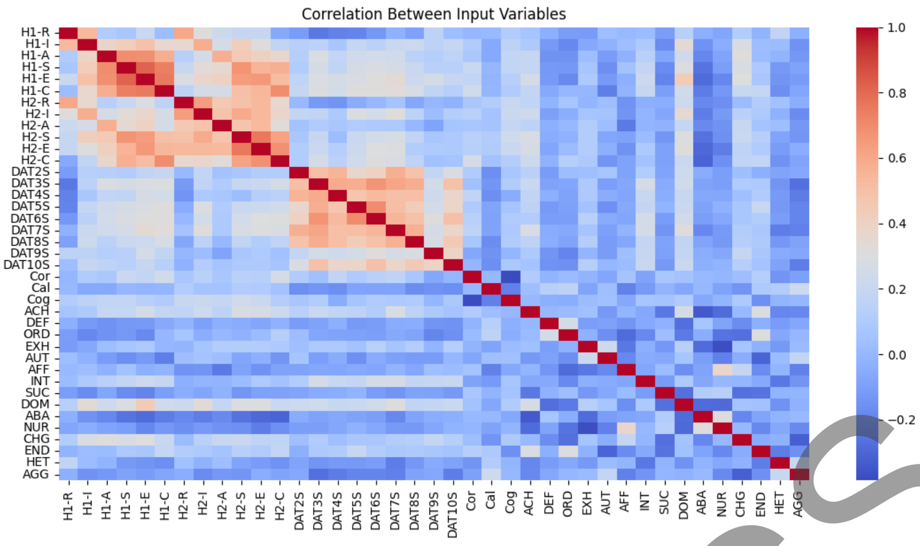

Fig. 5. Correlation between input variables.

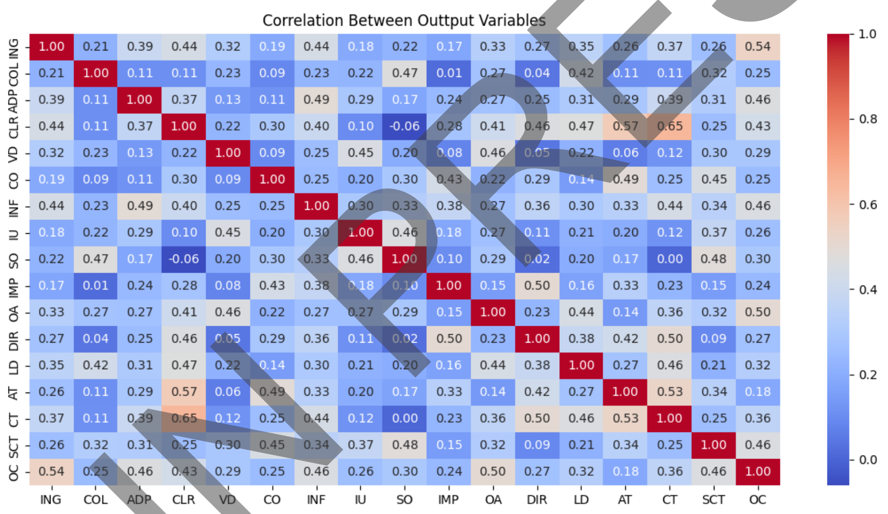

Fig. 6. Correlation between output variables.

ables, such as within the H1 and H2 categories, suggesting potential multicollinearity. Conversely, some variable pairs exhibit weak or negative correlations (indicated by shades of blue and white), suggesting a weaker linear association between them.

Next, the output variable is analyzed, as shown in Fig. 6. Based on that, the researchers can observe several moderate positive correlations (light red) between pairs of variables such as ING-COL, ING-ADP, COL-ADP, VD-AT, and LD-DIR (see the explanation of each variable in Table III). Conversely, there is a strong negative correlation (dark blue) between OC and several other variables, such as ING, COL, ADP, CLR, and VD. These correlation patterns indicate that certain competencies tend to develop together, while others exhibit contrasting behavioral tendencies. Such relationships provide important insights into how different competency dimensions interact within the psychological assessment framework. Understanding these patterns is useful for interpreting the structure of

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, “Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data”, CommIT Journal 20(1), 139–154, 2026.

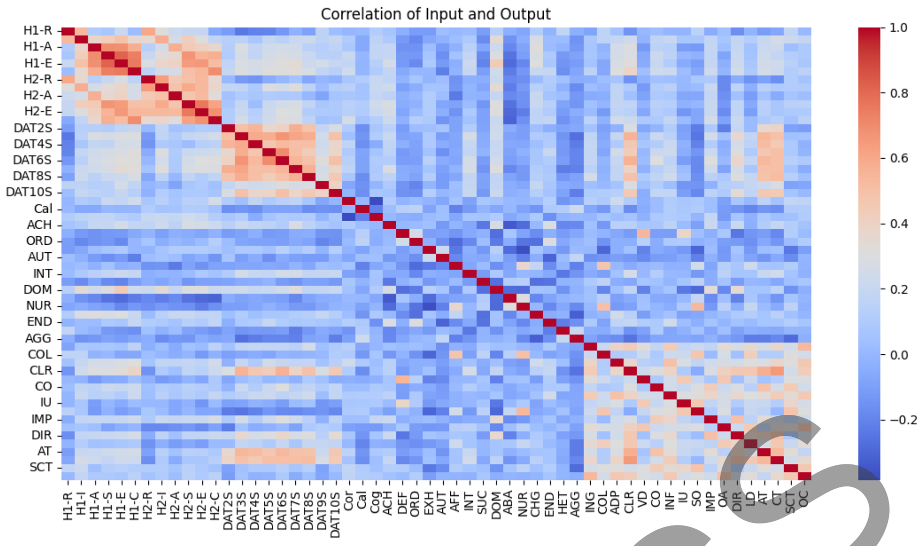

Fig. 7. Correlation of all variables (input and output).

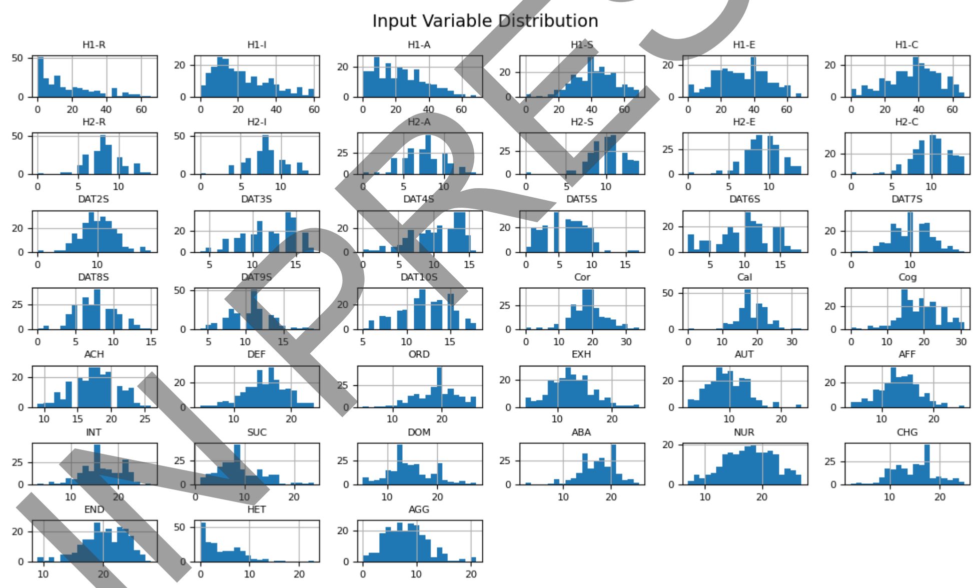

Fig. 8. Input variable distribution.

the competency model and supporting the development of more accurate predictive models.

To provide a comprehensive view, Fig. 7 presents a complete correlation analysis of all input and output variables. Warmer colors (reddish hues) indicate positive associations. In comparison, cooler colors (bluish hues) indicate negative associations, with color intensity reflecting the strength of the correlation with darker reds for stronger positive and darker blues for stronger negative relationships. Near-white colors suggest weak or negligible linear correlations. Fig. 8 shows the distribution of the input variables, illustrating how psychological assessment scores are spread across dimensions. The distributions vary in shape and spread. Some are roughly normal, while others are skewed or potentially multimodal.

The distribution of the output variables, representing competency levels derived from both psychological evaluations and specialized competency assessments, is illustrated in Fig. 9. The researchers observe that most

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, “Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data”, CommIT Journal 20(1), 139–154, 2026.

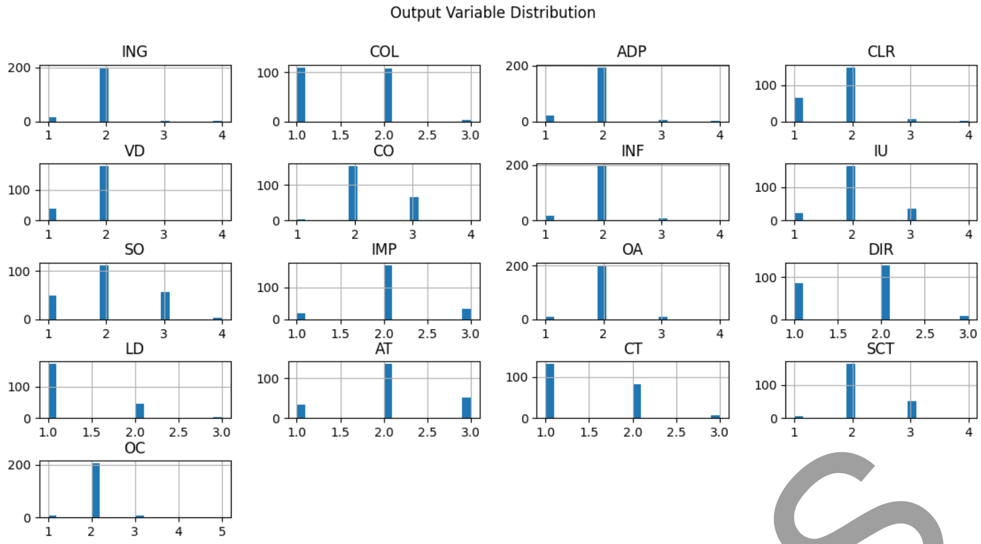

Fig. 9. Output variable distribution.

output variables have discrete, somewhat sparse distributions, often concentrated at specific integer values. These output values are distributed across predefined competency levels, ranging from 0 to 4, with each level representing a distinct degree of competency as assessed by standardized evaluation criteria.

C. Modelling

The training and testing process uses an 80:20 data split, with 80% for training and 20% for testing. The modeling phase involves applying several machine learning methods, as outlined in Table IV. Each model is trained and evaluated based on its performance, and hyperparameter tuning is conducted to optimize the results. After testing multiple models, the best-performing model is selected based on its accuracy and evaluation metrics. The final model is selected as the one that achieves the highest accuracy while maintaining a good balance between precision, recall, and overall generalization. The accuracy results for the tested models are shown in Table VI.

From Table VI, the Random Forest method achieves the highest accuracy of 81%. This result indicates that Random Forest is the most suitable method for this research context, as it outperforms other models in predictive performance. Consequently, this method is used as the final model for further analysis and implementation.

Before determining the best model, hyperparameter tuning is performed for each method to find the optimal parameters. For instance, in the Random Forest method, the tuning process involves adjusting parameters such as the number of trees, maximum depth, and minimum samples per split. The results of the hyperparameter tuning for Random Forest are shown in Fig. 10.

TABLE VI
PROPOSED METHOD ACCURACY.

Model	Mean Accuracy (%)
Random Forest	81.00
Gradient Boosting	78.44
Support Vector Machine	76.38
Neural Network	75.14
AdaBoost	77.84
K-Nearest Neighbors	75.18
Decision Tree	72.52

In the Random Forest method, the best parameters are 100 estimators and a maximum depth of 10. The number of estimators refers to the total number of decision trees used in the model. A higher number generally improves stability and accuracy although it also increases computational cost. Meanwhile, the maximum depth of 10 means that each Decision Tree in the Random Forest is limited to a depth of 10 levels. This limitation prevents overfitting by ensuring the model remains simple enough to generalize well to unseen data while still capturing important patterns. These parameter settings provide a balance between model complexity and generalization ability, ultimately leading to better accuracy in the research.

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, "Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data", CommIT Journal 20(1), 139–154, 2026.

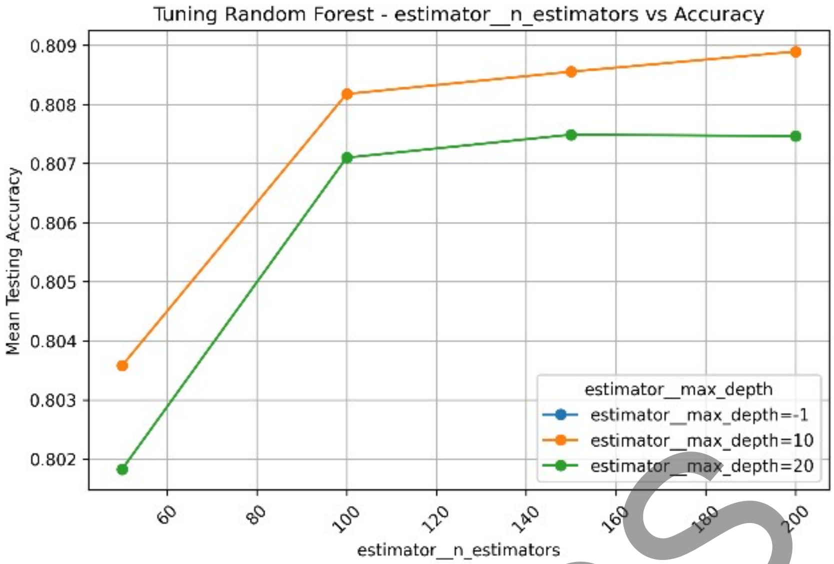

Fig. 10. Example of hyperparameter tuning for the Random Forest method.

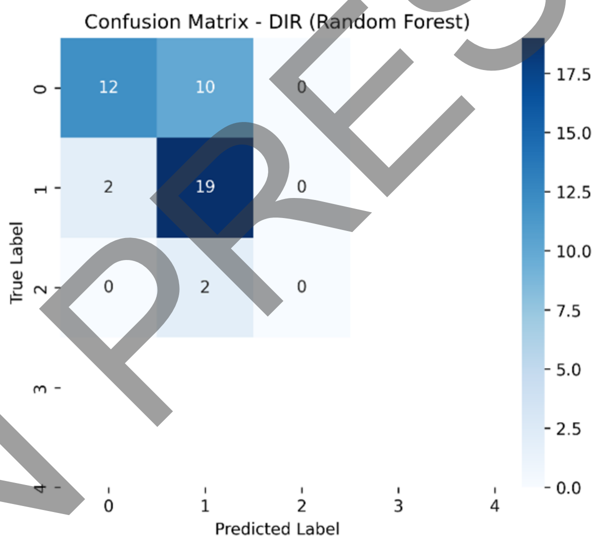

Fig. 11. Example of a confusion matrix for output competency - directing ability (DIR).

D. Model Result and Performance Evaluation

To evaluate the model, accuracy is calculated for each output (competency). The accuracy for each output is measured using precision, recall, and F1-Score. Precision measures how many of the predicted positive instances are actually correct, while recall measures how many of the actual positive instances are correctly identified. F1-Score represents the harmonic mean of precision and recall, providing a balanced assessment of the model's performance [51–53]. Additionally, a confusion matrix is used for each output to analyze the distribution of correct and incorrect predictions. The confusion matrix is shown in Fig. 11.

From these calculations, the accuracy of each output (competency) is determined. It represents the model's ability to predict each output using the applied multi-class classifier. A higher accuracy indicates that the model can effectively distinguish between different competency levels based on the given psychological assessment data. Additionally, the accuracy results for each output/competency are shown in Table VII. The results indicate that certain competencies can be predicted with high accuracy. Specifically, INF (91.11%), OA (95.56%), CT (95.56%), and OC (95.56%) achieve

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, “Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data”, CommIT Journal 20(1), 139–154, 2026.

TABLE VII
ACCURACY FOR EACH OUTPUT/COMPETENCY.

Output Variable	Accuracy (%)
COL	66.67
ADP	77.78
CLR	88.89
VD	84.44
CO	73.33
INF	91.11
IU	82.22
SO	48.89
IMP	77.78
OA	95.56
DIR	68.89
LD	84.44
AT	66.67
CT	95.56
SCT	66.67
OC	95.56
Mean	79.03

Note: the explanation of each competency can be seen in Table III.

accuracy above 90%, demonstrating that the model is highly effective in predicting these competencies. On the other hand, some competencies still exhibit lower prediction accuracy, such as COL (66.67%), DIR (68.89%), AT (66.67%), SCT (66.67%), and SO (48.89%). They have accuracy values below 70%. The explanation of each variable is in Table III. The results suggest that further improvements, such as feature engineering or adding additional input variables from other psychological assessments, may be required to enhance the model's predictive capability for these competencies. In other words, for certain competencies, the existing psychological assessment parameters may not be sufficient to improve prediction accuracy.

E. Interpretation/Explanation

The discussion section explains how to interpret the results and presents a new understanding of the problem after taking them into account. The researchers present only a few selected competencies, and the important predictive variables for those competencies are shown in the feature importance analysis. For example, the researchers focus on explaining the competencies of CT and OA.

As shown in Table V, the feature importance analysis indicates that the prediction of conceptual thinking (CT) competency is mainly driven by cognitive aptitude subtests, with the 7^{th} Differential Aptitude Test (DAT) subtest (DAT7S, importance = 0.113) as the most influential predictor. Other DAT subtests, DAT5S, DAT4S, DAT6S, and DAT8S, also make significant contributions, with importance values ranging from approximately 0.07 to 0.09. It reflects the strong connection between CT and abstract reasoning, logic, and problem-solving skills. Moderate influence is shown by DAT3S and DAT10S (importance 0.05–0.06). Personality variables, such as H1-I (Investigative) from the Holland model (importance = 0.0309) indicate that intellectual curiosity is associated with higher CT. Moreover, traits from the PAPI Kostick assessment, change (CHG) and orderliness (ORD) (importance = 0.016), suggest that adaptability and systematic behavior favor CT.

As shown in Table VIII, the feature importance analysis for predicting OA reveals that the most influential variable is the DAT2S subtest, emphasizing the role of specific cognitive abilities. Other DAT subtests, such as DAT4S, DAT5S, DAT7S, DAT8S, and DAT10S, also contribute significantly, highlighting the importance of various cognitive skills, including reasoning and problem-solving. Personality traits such as endurance (END), reflecting mental and physical stamina, and aggression (AGG), representing controlled assertiveness, play important roles in organizational contexts. Calculation ability (Cal) and orderliness (ORD) suggest that analytical skills and being systematic are valuable for managing organizational information effectively. Furthermore, Holland's personality types, such as Investigative (H2-I and H1-I), indicate that intellectual curiosity supports a deeper understanding of organizations. Meanwhile, Conventional (H1-C) reflects a preference for routine and structure, which are beneficial for navigating organizational roles. Social-related traits, including succorance (SUC) for the need for social support and Social Orientation (H2-S), influence interpersonal adaptation within organizations. Together, these cognitive and personality variables provide a comprehensive view of factors that predict strong OA.

F. Implication of the Study

The findings have both theoretical and practical implications. From a theoretical standpoint, the research contributes to the growing body of literature on the use of machine learning techniques in psychological assessment and human resource development. By demonstrating the feasibility of predicting employee competency levels from psychological data, the research provides evidence for integrating computational methods into traditional assessment frameworks. This integration can lead to more objective, data-driven approaches in understanding human behavior and potential. In the future, this model can also serve as a core component of AI-powered interview bots, enabling automated, real-time competency evaluation during virtual interviews, further streamlining the assessment process, and enhancing objectivity.

In practice, the developed model can be used by organizations to support talent management, recruitment,

Cite this article as: S. Suakanto, J. Siswanto, J. M. Pawlowski, M. Lubis, S. N. Lathifah, and L. W. Suwarsono, “Enhancing competency level prediction using machine learning: A data-driven approach based on psychological assessment data”, CommIT Journal 20(1), 139–154, 2026.

TABLE VIII
FEATURE IMPORTANCE FOR COMPETENCY ORGANIZATIONAL AWARENESS (OA).

Rank	Feature(s)	Importance Range	Interpretation
1	DAT2S	0.072	The 2nd Differential Aptitude Test (DAT) subtest is the most influential variable in predicting OA, likely related to cognitive abilities.
2–5	DAT4S, endurance (END), aggression (AGG), DAT5S	0.04–0.05	These features include cognitive subtests and personality traits such as endurance and aggression that influence OA.
6–10	DAT10S, calculation (Cal), DAT7S, orderliness (ORD), H2-I	0.03–0.04	Additional cognitive subtests and traits like calculation ability, orderliness, and investigative interest matter.
11–15	Succorance (SUC), H2-S, DAT8S, H1-C, H1-I	0.027–0.032	Personality and interest traits indicating social and cognitive engagement have moderate influence on OA.

promotion decisions, and training needs analysis. By predicting competencies from existing psychological assessment data, organizations can reduce reliance on costly, time-consuming manual evaluations. This approach also facilitates early identification of skill gaps and allows HR professionals to design personalized development programs for employees. Moreover, for large-scale organizations handling high volumes of assessment data, such models can serve as decision-support tools to improve efficiency and consistency in evaluating employee potential. Nevertheless, caution shall be taken in deploying such models operationally. Interpretability, fairness, and model generalizability need to be continuously evaluated, particularly when applying the model to new contexts or organizations with different cultural and demographic characteristics.

G. Strength of the Study

One of the main strengths of the research lies in the rigor of its methodology, which combines well-established psychological assessment techniques with advanced machine learning methods. By using real-world data from certified psychological evaluations and professional assessors, the research ensures high validity across its input and output variables. Careful data pre-processing, including handling outliers and missing values, enhances the reliability of the dataset used for modeling. Additionally, the research employs multiple machine learning algorithms and performs thorough hyperparameter tuning, enabling a comprehensive comparison and selection of the most effective model based on objective performance metrics. Another notable strength is the integration of SHapley Additive exPlanations (SHAP) analysis, which adds an interpretable layer to the model by identifying the most influential input features, thereby increasing transparency and trust in the machine learning outcomes. As a result, the research provides a robust, replicable framework that other researchers and practitioners can adapt or extend to similar contexts involving psychological assessment data and competency prediction.

IV. CONCLUSION

The results demonstrate that the Random Forest model achieves the highest accuracy of 81%, making it the most suitable method for competency prediction in this context. The model performs well in predicting certain competencies, particularly OA (95.56%), CT (95.56%), OC (95.56%), and INF (91.11%), indicating that the psychological assessment data effectively capture the patterns associated with these competencies. However, the model struggles with lower accuracy in predicting other competencies, suggesting that the current set of psychological assessment parameters may not be sufficient for these cases. Further enhancements, such as incorporating additional input variables from other assessment types, can improve the model’s predictive capability for these competencies. These findings highlight the potential of machine learning for competency prediction while underscoring the need for continuous improvement in feature selection and data augmentation to improve prediction accuracy across all competency areas.

From a practical perspective, these findings provide valuable insights for organizations seeking data-driven approaches to competency evaluation, enabling more efficient talent management and targeted employee development programs. Additionally, the research demonstrates the feasibility of integrating psychological assessment results with machine learning techniques to enhance decision-making in human resource practices.

However, the research has several limitations. The relatively small sample size of 221 respondents after data cleaning may limit the extent to which the findings can be applied to broader or more varied populations. The predictor variables are based solely on psychological assessment results, potentially excluding other important factors, such as job performance records, behavioural assessments, or peer feedback. In addition, the research is conducted within a single organizational setting, so its relevance to other sectors or cultural contexts has yet to be verified.

Future research can explore the integration of the proposed predictive model with LLM to enable more natural and context-aware interactions in competency assessment processes. For example, the model can be incorporated into an intelligent interview system or InterviewBot that dynamically generates questions and evaluates candidate responses in real time. In this scenario, psychological assessment results and conversational responses can be jointly analyzed to improve the accuracy and richness of competency predictions. Such integration will provide a more adaptive and scalable framework for automated competency evaluation in modern human resource management systems.

Acknowledgement

The authors acknowledge the financial support of the Ministry of Higher Education, Science, and Technology of Indonesia (Kemdiktisaintek) under Grant Number: 125/C3/DT.05.00/PL/2025, 7925/LL4/PG/2025, and 107/LIT07/PPM-LIT/2025.

Author Contribution

Designed the conceptual framework, S. S.; Developed the model, S. S.; Conducted implementation and testing, S. S.; Contributed from an HR perspective, assessed knowledge, J. S.; Assisted in model design, J. S.; Interpreted the results, J. S.; Provided insights from a competency perspective, J. M. P.; Contributed to HR considerations, J. M. P.; and Supported model design, J. M. P.; Conducted data exploration, M. L.; Defined relevant variables, M. L.; Collected required data, S. N. L.; Performed literature study, S. N. L.; Assisted in manuscript writing, S. N. L.; Contributed psychological perspectives, L. W. S.; Performed psychological assessment, L. W. S.; and Assisted in model design and interpreted results, L. W. S.

References

[1] S. Vaishnavi and G. Rajini, "Prediction of competencies during selection process: Correlation and mediation analysis," *Russian Law Journal*, vol. 11, no. 3, pp. 50–66, 2023.

[2] S. Kadam, M. Negi, D. Khairnar, S. Nalawade, and S. Patil, "Student's competency level prediction," *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 9, no. VI, pp. 4478–4480, 2023.

[3] C. Cao and Z. Zhang, "Machine learning-assisted competency modeling for human resource management jobs," *Mobile Information Systems*, vol. 2022, no. 1, pp. 1–15, 2022.

[4] K. Agustian, A. Pohan, A. Zen, W. Wiwin, and A. J. Malik, "Human resource management strategies in achieving competitive advantage in business administration," *Journal of Contemporary Administration and Management (ADMAN)*, vol. 1, no. 2, pp. 108–117, 2023.

[5] B. Škrinjarić, "Competence-based approaches in organizational and individual context," *Humanities and social sciences communications*, vol. 9, no. 1, pp. 1–12, 2022.

[6] V. A. Jerin, "Potential appraisal and development in Human Resource Development (HRD): A review," *Asian Journal of Agricultural Extension, Economics & Sociology*, vol. 39, no. 2, pp. 111–117, 2021.

[7] D. Y. Dai, "Rethinking human potential from a talent development perspective," *Journal for the Education of the Gifted*, vol. 43, no. 1, pp. 19–37, 2020.

[8] J. M. Conte and R. K. Harmata, "Person-centered study of cognitive ability dimensions using latent profile analysis," *Journal of Intelligence*, vol. 11, no. 5, pp. 1–12, 2023.

[9] P. B. Andreatta, C. H. Renninger, M. W. Bowyer, and J. M. Gurney, "Measuring competency: Improving the validity of your procedural performance assessments," *Annals of Surgery Open*, vol. 4, no. 4, pp. 1–7, 2023.

[10] W. Febri Surya and A. Kasim, "The role of assessment center in career development in POLDA Bangka Belitung Islands," *International Review of Humanities Studies*, vol. 7, no. 2, pp. 457–472, 2022.

[11] C. Atkinson, J. Barrow, and S. Norris, "Assessment practices of educational psychologists and other educational professionals," *Educational Psychology in Practice*, vol. 38, no. 4, pp. 347–363, 2022.

[12] T. V. Tran, S. Lepistö, and J. Järvinen, "The relationship between subjectivity in managerial performance evaluation and the three dimensions of justice perception," *Journal of Management Control*, vol. 32, no. 3, pp. 369–399, 2021.

[13] J. M. Cucina, N. R. Martin, N. L. Vasilopoulos, and H. F. Thibodeuax, "Self-serving bias effects on job analysis ratings," *The Journal of Psychology*, vol. 146, no. 5, pp. 511–531, 2012.

[14] K. Y. Wilson, "An analysis of bias in supervisor narrative comments in performance appraisal," *Human Relations*, vol. 63, no. 12, pp. 1903–1933, 2010.

[15] A. Alsharif, "Transforming HR Practices: Using AI to revolutionize employee performance management," *International Journal of Academic Publishing in Educational Sciences and Humanities (IJAPESH)*, pp. 43–51, 2025.

[16] R. N. Landers, E. M. Auer, L. Dunk, M. Langer, and K. N. Tran, "A simulation of the impacts of machine learning to combine psychometric employee selection system predictors on performance prediction, adverse impact, and number of dropped predictors," *Personnel Psychology*, vol. 76, no. 4, pp. 1037–1060, 2023.

[17] Y. Liu, "Application of machine learning in human resources management: Research on employee performance prediction model," in *Proceedings of the International Conference on Decision Science & Management*. Hong Kong: Association for Computing Machinery, April 26–28, 2024, pp. 134–139.

[18] Z. Nayem and M. A. Uddin, "Unbiased employee performance evaluation using machine learning," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 10, no. 1, pp. 1–11, 2024.

[19] N. Talpur, N. C. G. Alier, H. Matsom, M. H. A. Abdullah, and S. Khatoon, "Predicting employee performance using machine learning to enhance workforce efficiency," in *2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*. Pune, India: IEEE, Aug. 23–24, 2024, pp. 1–4.

[20] J. Huang, L. Li, L. Zhang, X. Qiao, and F. You, "Exploration of a deep learning-based mechanism for predicting the work competence of community caregivers," *Applied Mathematics and Nonlinear Sciences*, vol. 9, no. 1, pp. 1–14, 2024.

[21] Z. Wei and L. Shuquan, "Prediction model of safety competency of construction workers based on machine learning," *China Safety Science Journal*, vol. 33, no. 7, pp. 51–57, 2023.

[22] G. Ibarra-Vazquez, M. S. Ramírez-Montoya, M. Buenestado-Fernández, and G. Olague, "Predicting open education competency level: A machine learning approach," *Heliyon*, vol. 9, no. 11, pp. 1–15, 2023.

[23] R. Farooqi and N. Iqbal, "Performance evaluation for competency of bank telemarketing prediction using data mining techniques," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 5666–5674, 2019.

[24] S. Fukui, W. Wu, J. Greenfield, M. P. Salyers, G. Morse, J. Garabrant, E. Bass, E. Kyere, and N. Dell, "Machine learning with human resources data: Predicting turnover among community mental health center employees," *Journal of Mental Health Policy and Economics*, vol. 26, no. 2, pp. 63–76, 2023.

[25] F. Guerranti and G. M. Dimitri, "A comparison of machine learning approaches for predicting employee attrition," *Applied Sciences*, vol. 13, no. 1, pp. 1–8, 2022.

[26] M. Fadel, K. Kanasfi, and A. Wibowo, "Application of machine learning in predicting employee discipline violations in financial service company," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 1, pp. 171–178, 2024.

[27] C. Zhu, H. Zhu, H. Xiong, C. Ma, F. Xie, P. Ding, and P. Li, "Person-job fit: Adapting the right talent for the right job with joint representation learning," *ACM Transactions on Management Information Systems (TMIS)*, vol. 9, no. 3, pp. 1–17, 2018.

[28] H. Gülten and H. Baraçlı, "A machine learning-based forecast model for career planning in human resource management: A case study of the Turkish post corporation," *Applied Sciences*, vol. 14, no. 15, pp. 1–20, 2024.

[29] N. Dawson, M.-A. Rizoiu, B. Johnston, and M.-A. Williams, "Predicting skill shortages in labor markets: A machine learning approach," in *2020 IEEE International Conference on Big Data (Big Data)*. Atlanta, GA, USA: IEEE, Dec. 10–13, 2020, pp. 3052–3061.

[30] K. G. Ayu, D. W. Sari, I. Farida, D. Harsono, R. W. Kosaman, I. H. Sumitro, F. N. Iman, Mansuri, E. R. Kaburuan, and M. A. Fitri, "Classification of employee competency assessment using Naïve Bayes and K-Nearest neighbor (KNN) algorithms," *Journal of Advances in Information Technology*, vol. 15, no. 7, pp. 879–885, 2024.

[31] M. Harirchian, F. Amin, S. Rouhani, A. Aligholipour, and V. A. Lord, "AI-enabled exploration of Instagram profiles predicts soft skills and personality traits to empower hiring decisions," 2022. [Online]. Available: https://arxiv.org/abs/2212.07069

[32] J. Park, Y. Feng, and S. P. Jeong, "Developing an advanced prediction model for new employee turnover intention utilizing machine learning techniques," *Scientific Reports*, vol. 14, no. 1, pp. 1–14, 2024.

[33] Y. Choi and J. W. Choi, "A study of job involvement prediction using machine learning technique," *International Journal of Organizational Analysis*, vol. 29, no. 3, pp. 788–800, 2021.

[34] C. Chaiwichit and S. Samart, "The empirical research of teamwork competency factors and prediction on academic achievement using machine learning for students in thailand," *Perspectives of Sciences & Education*, vol. 57, no. 3, pp. 540–556, 2022.

[35] K. Alipour, A. Ray, X. Lin, J. P. Schulze, Y. Yao, and G. T. Burachas, “The impact of explanations on AI competency prediction in VQA,” in *2020 IEEE International Conference on Humanized Computing and Communication with Artificial Intelligence (HCCAI)*. Irvine, CA, USA: IEEE, Sep. 21–23, 2020, pp. 25–32.

[36] M. Şimşek and A. S. Daş, “The effect of handling imbalanced datasets methods on prediction of entrepreneurial competency in university students,” *Turkish Journal of Forecasting*, vol. 6, no. 2, pp. 53–60, 2022.

[37] P. Arjyotha, V. Promasatayaprot, and T. Promaruk, “Competency prediction equation for the performance of academic health personnel in subdistrict health promoting hospitals in the northeastern region of Thailand,” *Journal of Southwest Jiaotong University*, vol. 56, no. 2, pp. 1–11, 2021.

[38] M. F. N. Adillah, S. Suakanto, and N. I. Utama, “Implementation of machine learning-based classification model in employee recruitment decision prediction,” *Journal La Multiapp*, vol. 6, no. 2, pp. 341–352, 2025.

[39] L. M. Spencer and S. M. Spencer, “Competence at work: Models for superior performance,” 1993.

[40] J. M. A. Pacheco, “Gestión por competencias: Propuesta de cara a la realidad organizacional en Perú (Competencies management: Proposal facing the organizational reality in Peru),” *Revista de Filosofia (Venezuela)*, no. 98, pp. 310–326, 2021.

[41] R. Nurahaju and N. S. Widanti, “Effect of competencies and personality on employee performance,” *Research on Humanities and Social Sciences*, vol. 10, no. 24, pp. 53–58, 2020.

[42] D. P. Diwanti and M. Hariyanto, “The influence of human resource competence in Spencer’s concept to organizational entrepreneurship,” *Jurnal Manajerial*, vol. 9, no. 02, pp. 117–137, 2022.

[43] A. Barman and K. Das, “Whether B-Schools Care Spencer & Spencer’s workplace competency framework in the 21st century? Revalidating through reliability,” *International Journal of Advanced Science and Technology*, vol. 29, no. 11, pp. 2910–1920, 2020.

[44] Y. Zou and C. Gao, “Extreme learning machine enhanced gradient boosting for credit scoring,” *Algorithms*, vol. 15, no. 5, pp. 1–20, 2022.

[45] R. Genuer, “Variance reduction in purely random forests,” *Journal of Nonparametric Statistics*, vol. 24, no. 3, pp. 543–562, 2012.

[46] D. Valkenborg, A. J. Rousseau, M. Geubbelmans, and T. Burzykowski, “Support vector machines,” pp. 754–757, 2023.

[47] G. Li, M. Müller, B. Ghanem, and V. Koltun, “Training graph neural networks with 1000 layers,” in *Proceedings of the 38th International Conference on Machine Learning*. PMLR, 2021, pp. 6437–6449.

[48] M. Açıkkar and S. Tokgöz, “An improved KNN classifier based on a novel weighted voting function and adaptive k-value selection,” *Neural Computing and Applications*, vol. 36, no. 8, pp. 4027–4045, 2024.

[49] M. Zaman and A. Hassan, “Fuzzy heuristics and decision tree for classification of statistical feature-based control chart patterns,” *Symmetry*, vol. 13, no. 1, pp. 1–12, 2021.

[50] O. Hornyák and L. B. Iantovics, “Adaboost algorithm could lead to weak results for data with certain characteristics,” *Mathematics*, vol. 11, no. 8, pp. 1–24, 2023.

[51] B. A. Hamed, O. A. S. Ibrahim, and T. Abd El-Hafeez, “Optimizing classification efficiency with machine learning techniques for pattern matching,” *Journal of Big Data*, vol. 10, no. 1, pp. 1–18, 2023.

[52] S. A. Hicks, I. Strümke, V. Thambawita, M. Hammou, M. A. Riegler, P. Halvorsen, and S. Parasa, “On evaluation metrics for medical applications of artificial intelligence,” *Scientific Reports*, vol. 12, pp. 1–9, 2022.

[53] J. Somasekar, B. Unhelkar, S. S. S., and P. Chakrabarti, “A novel approach to movie recommendation using weighted collaborative filtering with activity and rating variability analysis,” *Journal of Informatics Education and Research*, vol. 4, no. 3, pp. 2313–2322, 2024.