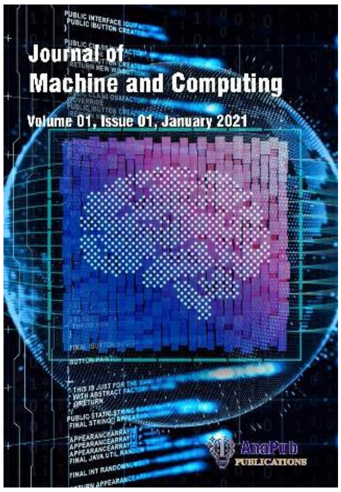
Journal Pre-proof



Enhancing Strategy and Governance Through AI-Driven Behavioral Competency Analytics: An ML Model for Competency Development

Srinivasa Rao Dasaraju, Venkata Raghu Babu Nallamalli, Jayanthi Rajendran, Madhusudhana Rao Chennamsetty, Vipin Jain and Girish Kumar Painoli

DOI: 10.53759/7669/jmc202505198

Reference: JMC202505198

Journal: Journal of Machine and Computing.

Received 02 January 2025

Revised from 10 May 2025

Accepted 07 August 2025

Please cite this article as: Srinivasa Rao Dasaraju, Venkata Raghu Babu Nallamalli, Jayanthi Rajendran, Madhusudhana Rao Chennamsetty, Vipin Jain and Girish Kumar Painoli, "Enhancing Strategy and Governance Through AI-Driven Behavioral Competency Analytics: An ML Model for Competency Development", Journal of Machine and Computing. (2025). Doi: https:// doi.org/10.53759/7669/jmc202505198.

Enhancing Strategy and Governance Through AI-Driven Behavioral Competency Analytics: An ML Model for Competency Development

Srinivasa Rao Dasaraju[1], Venkata Raghu Babu Nallamalli[2], Jayanthi Rajendran[3,*], Madhusudhana Rao Chennamsetty[4], Vipin Jain[5], Girish Kumar Painoli[6]

[1]Finance and Accounting, IBS Hyderabad (Under IFHE, Hyderabad), Telangana, 501203, India. *Email: srinivasa.rao@ibsindia.org*

[2]Department of Management Studies, RISE Krishna Sai Prakasam Group of Institutions, Andhra Pradesh, 523272, India. *E-mail: dr.nvraghubabu@gmail.com*

[3]Department of English, Easwari Engineering College, Chennai, 6000[illegible], Tamil Nadu, India.
**Corresponding Author E-mail: jayanthirajendran@ee[illegible]srmrm[illegible]edu.in*

[4]School of Computing, Mohan Babu University, Tirupati, Andhra Pradesh, 517102, India. *Email: npr4567@gmail.com*

[5]Department of Management, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, 244102, India. *Email: vipin555@rediffmail.com*

[6]University Institute of Management and Commerce, Guru Nanak University, Hyderabad, Telangana, 501506, India. *Email: gkpainoli@gmail.com*

Abstract

Strategic decision-making and organizational governance increasingly depend on accurate assessment of human behavioral competencies. Traditional evaluation methods often lack scalability, objectivity, and predictive insight, limiting their utility in dynamic enterprise environments. This study proposes a machine learning-based framework for competency development and analytics that integrates multi-source behavioral data with predictive modeling to enable data-driven governance. A structured pipeline is developed comprising behavioral signal alignment, feature engineering, probabilistic classification, and governance-aligned scoring. The framework is operationalized using multiple supervised learning models, including Logistic Regression, Random Forest, XGBoost, and Multilayer Perceptron, with XGBoost achieving the highest classification accuracy (83.4%) and superior probabilistic calibration. Cross-validation confirmed the robustness of performance with minimal variance (±1.5%), and interpretability was supported through feature attribution. Behavioral profiling revealed high central tendency in Analytical Thinking and wide dispersion in Ethical Conduct, informing strategic prioritization. The proposed model delivers calibrated, interpretable, and governance-compatible competency predictions, presenting a scalable solution for institutional leadership development, risk management, and policy alignment. Experimental validation

across 1,247 behavioral instances confirms the model's effectiveness in bridging human capital analytics with strategic decision processes.

Keywords: Behavioral Competency, Machine Learning, XGBoost, Strategic Governance, Competency Profiling, Probabilistic Calibration, Human Capital Analytics

1. Introduction

Organizational performance in contemporary knowledge economies is increasingly determined by the behavioral competencies of individuals rather than solely by technical capabilities or domain expertise [1]. As enterprises adapt to rapidly shifting market conditions, strategic priorities such as leadership effectiveness, adaptability, ethical conduct, and cognitive agility have become essential drivers of sustained success [2]. These competencies impact not only internal operational cohesion but also external stakeholder confidence, regulatory compliance, and long-term innovation potential. Consequently, the measurement, development, and deployment of behavioral competencies have emerged as critical components of strategic governance and workforce transformation [3].

Despite their importance, conventional competency assessment practices—such as structured interviews, supervisor evaluations, and self-assessment inventories—are limited by subjectivity, evaluator bias, low scalability, and insufficient integration with real-time decision systems [4]. These methods typically propose static, retrospective snapshots of employee performance, lacking the predictive granularity required for high-stakes decisions related to leadership succession planning, organizational risk profiling, and regulatory alignment. As a result, organizations face a growing imperative to adopt objective, data-driven approaches that can systematically evaluate behavioral attributes across large, diverse populations while preserving interpretability and decision accountability [5].

Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) propose transformative opportunities for competency analytics. Supervised learning algorithms, in particular, are capable of mapping complex behavioral features to latent competency classes using both structured and unstructured data sources, such as psychometric assessments, communication patterns, 360-degree feedback, and HR information systems [6, 7]. When properly calibrated, these models can deliver probabilistic predictions with quantifiable confidence levels, enabling downstream applications in decision support, performance management, and leadership development. However, the adoption of such systems for governance purposes demands rigorous attention to fairness, reliability, transparency, and actionable interpretability—criteria often unmet by black-box AI solutions [8].

This research addresses these challenges by proposing a comprehensive, explainable, and governance-compatible model for behavioral competency analytics grounded in ML. The model integrates multi-source behavioral signals into an engineered feature space, employs supervised classification models to infer competency classes, and generates probabilistic outputs used to compute governance-aligned scores. Model development follows best practices in cross-validation, calibration testing, and interpretability auditing to ensure the integrity and utility of predictions.

The study contributes to the literature by formalizing a competency modeling pipeline that aligns technical rigor with strategic relevance. Unlike prior efforts that focus narrowly on performance classification or psychometric diagnostics, the proposed approach is holistic, linking individual-level behavioral insights to macro-level governance objectives. It further evaluates model performance not only through standard accuracy metrics but also through probabilistic calibration measures and class-wise behavioral profiling, ensuring the robust and responsible deployment of AI in human capital management contexts.

The remainder of this paper is organized as follows: Section 2 reviews the existing literature on behavioral competency models and AI applications in workforce analytics. Section 3 describes the methodology, including data preprocessing, Feature Engineering (FE), model training, and evaluation design. Section 4 presents empirical results from the classification, calibration, and interpretability analysis. Section 5 concludes with future research directions and considerations for deployment.

2. Literature Review

The integration of AI into human resource management and governance models has accelerated the development of advanced systems for competency identification, workforce planning, and Strategic Decision-Making (SDM). This section reviews prior work relevant to AI-driven talent analytics, behavioral competency modeling, and ML employed in predictive evaluation systems. The review is structured around three thematic pillars: (1) AI in talent analytics and workforce systems, (2) competency modeling and behavioral measurement, and (3) ML for performance prediction and interpretability.

2.1 AI in Talent Analytics and Strategic Governance

The emergence of AI as a catalyst for workforce transformation has sparked growing interest in intelligent talent analytics systems that can extract actionable insights from behavioral and organizational data.

[9] Provide a comprehensive survey of AI techniques applied to talent analytics, identifying core components such as data fusion, behavioral FE, model calibration, and

decision support integration. The study categorizes AI tools into predictive, prescriptive, and adaptive analytics models, emphasizing the importance of transparency and explainability, particularly in applications that impact promotion, compensation, and succession planning. Their taxonomy establishes the theoretical foundation for integrating AI outputs into governance workflows, where decisions must align with fairness and accountability standards.

Similarly, [10] explored technology acceptance through a behavioral lens using ML models applied to fintech transaction data. Their study validates the use of decision trees and gradient boosting in modeling latent behavioral responses and confirms the effectiveness of probabilistic classifiers in capturing digital interaction patterns. These insights support the relevance of ML-based behavioral inference systems in broader domains beyond fintech, including education, human capital management, and competency development.

2.2 Competency Modeling and Behavioral Structuring

The transition from traditional competency assessments to digital, AI-augmented systems requires formal models for defining, measuring, and validating behavioral indicators. [11] proposed the Meta AI Literacy Scale (MAILS), a structured instrument for evaluating AI-related competencies across cognitive, emotional, and strategic dimensions. Their model introduces meta-competency categories such as self-regulation and situational awareness, which closely align with enterprise-level governance objectives. By grounding competency definitions in psychological theory and empirical testing, MAILS facilitates the transformation of abstract behavioral traits into quantifiable model features.

[12] further expanded the competency modeling literature by proposing a hierarchical model for AI literacy, rooted in constructivist theory and validated through iterative expert consultations. Their approach formalizes the competency lifecycle—from conceptual model to measurable indicators—and outlines a roadmap for integrating assessment metrics into educational and professional development systems. Both works reinforce the importance of structured, theory-informed competency definitions when designing AI-driven classification and scoring systems.

In an applied context, [13] examined teaching competencies in higher education under the influence of AI integration. Their findings revealed a multidimensional competency model encompassing technical fluency, communication, ethical reasoning, and instructional adaptability. The study provides empirical validation of how AI exposure reshapes expected behavioral attributes and proposes a practical basis for model training datasets that incorporate domain-specific competency clusters.

2.3 ML for Behavioral Prediction and Interpretability

ML proposals are powerful tools for modeling non-linear relationships between behavioral inputs and competency outcomes.

[14] Conducted a scientometric and empirical analysis on behavior-driven learning performance prediction. The study compared models such as XGBoost, Random Forest (RF), and neural networks, and identified XGBoost as the most stable and interpretable classifier when paired with SHAP (SHapley Additive exPlanations) for feature attribution. Their findings confirm the suitability of ensemble-based methods for modeling behavioral systems that demand both predictive strength and decision transparency.

Supporting this, [15] demonstrated the efficacy of XGBoost in predicting educational performance, outperforming baseline models in both accuracy and reliability. Their study emphasized the importance of feature selection, class balance, and probabilistic calibration in achieving meaningful results. The use of interpretable outputs, including reliability plots and class-wise scoring, further bridges the gap between algorithmic outputs and stakeholder comprehension—a necessary feature in governance applications.

2.4. Summary and Research Gap

Existing research establishes a robust foundation for the application of AI in behavioral competency modeling. Prior studies provide theoretical competency taxonomies, validated scoring instruments, and empirical support for ensemble learning and explainable models. However, a critical gap remains in the **end-to-end operationalization** of behavioral competency analytics models that integrate engineered behavioral signals, probabilistic ML outputs, and governance-aligned scoring mechanisms. Moreover, few studies simultaneously address model calibration, feature interpretability, and domain-specific profiling within a single architecture. This research addresses these gaps by developing a calibrated, explainable, and governance-compatible ML pipeline for behavioral competency evaluation using multi-source data and interpretable modeling techniques [16-19].

3. Methodology

A rigorous methodology is essential to operationalize behavioral competency analytics within strategic and governance systems. The proposed methodology integrates ML with behavioral data mining to establish a systematic, scalable, and explainable model for competency evaluation. This section outlines the conceptual foundation, data flow architecture, modeling techniques, and evaluation protocols adopted in the construction of the AI-driven competency analytics model.

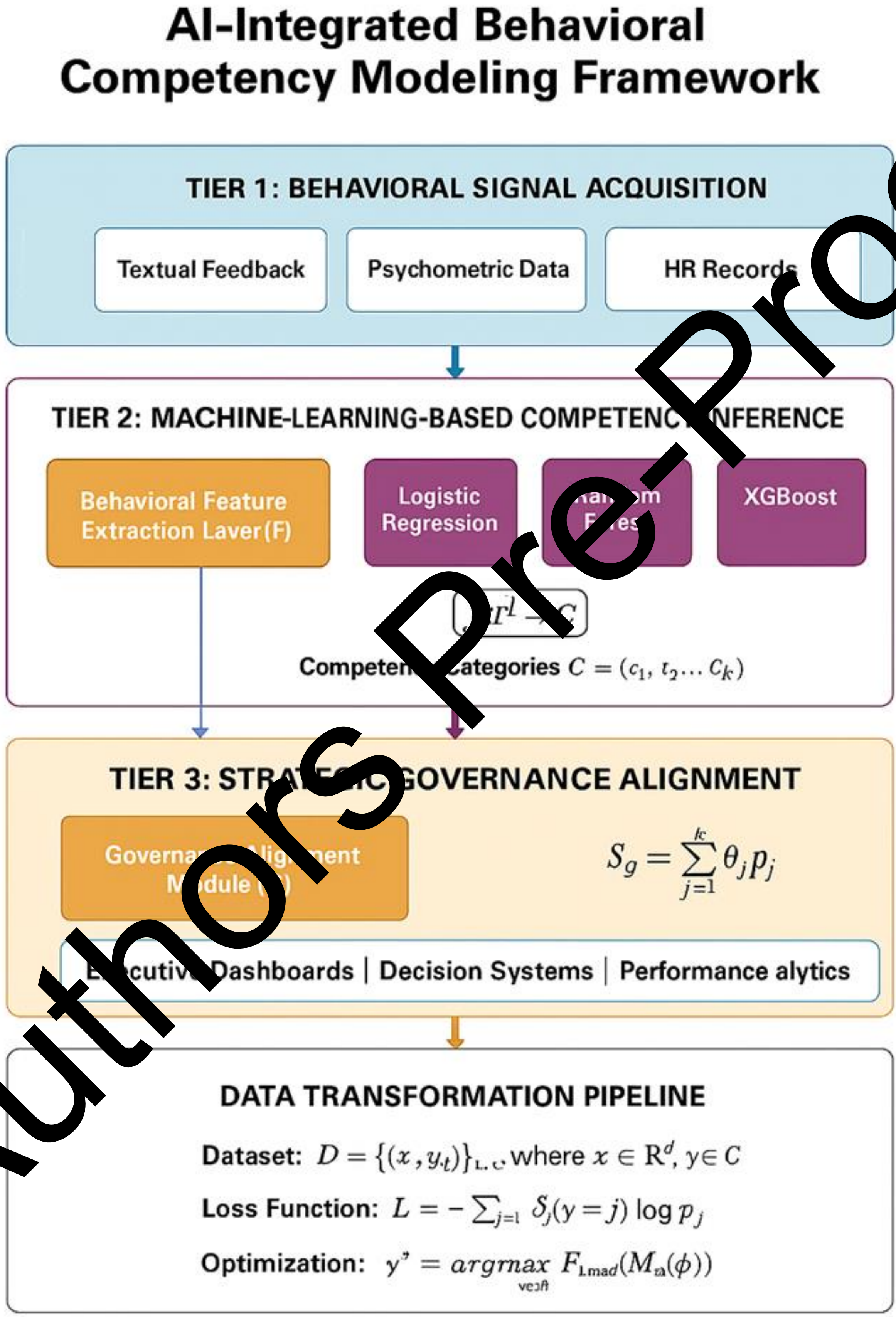


Figure 1: Conceptual Model

3.1 Conceptual Model

The conceptual model (Figure 1) establishes the theoretical and architectural basis for integrating AI into behavioral competency modeling, linking individual-level attributes to broader strategic governance outcomes. This integration is facilitated by a three-tiered system encompassing behavioral signal acquisition, ML-based competency inference, and strategic governance alignment.

Let the dataset denote a behavioral observation space.

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N} \quad (1)$$

where $x_i \in \mathbb{R}^d$ represents the i-th individual's feature vector consisting of d behavioral indicators, and $y_i \in \mathcal{C}$ is the corresponding competency class label from a predefined set of competency categories $\mathcal{C} = \{c_1, c_2, \dots, c_k\}$. Here, N denotes the total number of observed individuals in the dataset. The objective is to learn a function.

$$f: \mathbb{R}^d \to \mathcal{C} \quad (2)$$

that maps each feature vector x_i to its predicted competency class $\hat{y}_i = f(x_i)$, enabling automated classification of behavioral profiles.

To evaluate strategic alignment, a governance impact score is computed using the derived competencies. Let $\theta_j \in \mathbb{R}$ denote the impact weight associated with the competency category c_j, and let p_j denote the predicted probability that an individual belongs to a competency c_j. The aggregate governance alignment score S_g is given by:

$$S_g = \sum_{j=1}^{k} \theta_j \cdot p_j \quad (3)$$

where $S_g \in \mathbb{R}$ represents a scalar index capturing the strategic value contribution of a behavioral profile. The values of p_j are obtained from the softmax outputs of the trained model, while the weights θ_j are determined through expert elicitation or regression modeling linking competencies to organizational performance indicators.

To structure the data transformation pipeline, the entire competency evaluation architecture is decomposed into three primary functional modules:

1. **Behavioral Feature Extraction Layer ($\mathcal{F}$) :** Transforms raw inputs (e.g., textual feedback, psychometrics, HR data) into standardized feature vectors x_i using natural language processing, signal aggregation, or embedding functions.
2. **Competency Inference Engine ($\mathcal{M}$):** Implements the learned mapping $f(\cdot)$ via supervised ML (*e.g.,* RF, SVM, neural networks), producing class predictions $\hat{y}_i$ and probability vectors $[p_1, \dots, p_k]$.

3. **Governance Alignment Module ($\mathcal{G}$) :** Computes the final governance score S_g based on equation (3), enabling integration of competency analytics into executive dashboards and decision systems.

The final output of the model is a structured mapping:

$$\mathcal{D} \xrightarrow{\mathcal{F}} \mathbb{R}^d \xrightarrow{\mathcal{M}} \mathcal{C} \xrightarrow{\mathcal{G}} \mathbb{R} \tag{4}$$

This end-to-end transformation facilitates data-driven SDM based on objective behavioral analytics.

The conceptual models thus form the foundation for a scalable and explainable competency analytics system that bridges individual behavioral data with organizational governance insights. Subsequent sections describe the data modeling procedures, ML employed, and the system's empirical validation.

3.2 Data Collection and Preprocessing

Accurate and high-quality data acquisition forms the foundation for any ML-based behavioral competency analysis. This section describes the sources, structure, and preprocessing protocols applied to the behavioral datasets used for competency inference. Emphasis is placed on ensuring standardization, ethical compliance, and consistency throughout the transformation process to enable reliable model training and interpretation.

3.2.1 Behavioral Data Sources

The behavioral data used for competency modeling were drawn from a diverse set of organizational repositories, each contributing a specific dimension of behavioral expression:

- **Performance Appraisal Reports**: Structured annual feedback forms containing ratings on soft skills, communication style, adaptability, and teamwork.
- **360-Degree Feedback**: Multi-source evaluations collected from supervisors, peers, and subordinates, covering dimensions of leadership, conflict resolution, and ethical conduct.
- **Digital Communication Logs**: Linguistic and sentiment features extracted from corporate emails, meeting transcripts, and internal messaging platforms.
- **Psychometric Assessments**: Standardized test scores reflecting traits such as openness, conscientiousness, and emotional stability.
- **HRIS Metadata**: Demographic attributes, promotion timelines, and tenure records, used for auxiliary features and stratification.

These multi-source inputs contribute to the generation of a unified behavioral profile vector $x_i \in \mathbb{R}^d$ as defined in Equation (1).

3.2.2 Data Cleaning and Anonymization

Raw data collected from multiple systems often contains inconsistencies, missing values, and identifying information. A formal cleaning process was applied:

- **Imputation:** Missing values were filled using a hybrid approach that combined statistical mean imputation for numeric fields with the mode for categorical variables, ensuring statistical consistency while minimizing data leakage.
- **Deduplication:** Records with identical identifiers and timestamp overlaps were removed to avoid redundancy.
- **Normalization:** All numerical features were scaled using min-max normalization as:

$$x_{ij}^{\text{norm}} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (5)$$

  where x_{ij} is the original value of the j-th feature for the i-th individual, and $\min(x_j), \max(x_j)$ denote the minimum and maximum values of feature j across the dataset. This maps all values into the $[0, 1]$ range, preserving scale invariance across features.
- **Anonymization:** Personally identifiable information (PII) was removed or tokenized to ensure ethical data handling and compliance with relevant regulations. Unique IDs were assigned to each participant using a cryptographic hash function.

3.2.3 Feature Vector Construction

After standardization, behavioral indicators were aggregated into a structured feature matrix:

$$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{bmatrix} \in \mathbb{R}^{N \times d} \quad (6)$$

where each row vector $x_i \in \mathbb{R}^d$ represents the cleaned, normalized behavioral profile of the i-th individual and each column corresponds to a specific behavioral or psychometric attribute. The matrix X serves as the model input for the ML engine described in subsequent sections.

3.2.4 Label Encoding and Class Balancing

Competency labels $y_i \in \mathcal{C}$ were encoded using ordinal or categorical schemes depending on the model design. In cases of imbalanced class distribution, Synthetic Minority Over-sampling Technique (SMOTE) was applied to augment underrepresented classes, ensuring adequate representation during model training without distorting feature semantics.

3.3 Feature Engineering

FE is a critical methodological step that transforms raw behavioral inputs into high-dimensional, discriminative representations suitable for ML-based competency inference. This section outlines the design of domain-relevant behavioral features, the transformation of heterogeneous input types, and the dimensional reduction strategies employed to optimize model performance while maintaining interpretability.

3.3.1 Behavioral Feature Taxonomy

The behavioral features were classified into four functional categories, each capturing a distinct aspect of individual workplace behavior:

- **Linguistic Features ($\mathcal{F}_1$) :** Extracted from textual sources such as emails and performance narratives using natural language processing (NLP). These include word frequency vectors, syntactic complexity, tone polarity, and sentiment scores.
- **Interactional Features ($\mathcal{F}_2$) :** Derived from communication metadata including message response latency, participation in collaborative platforms, and meeting contribution frequency.
- **Psychometric Features ($\mathcal{F}_3$) :** Numerical variables obtained from standardized assessments capturing personality traits, cognitive agility, and emotional intelligence metrics.
- **Historical and Structural Features ($\mathcal{F}_4$) :** Attributes reflecting career progression, tenure, department, and previous role transitions.

The complete feature vector for an individual x_i is structured as:

$$x_i = [\mathcal{F}_1(i), \mathcal{F}_2(i), \mathcal{F}_3(i), \mathcal{F}_4(i)] \in \mathbb{R}^d \quad (7)$$

where x_i is the concatenation of sub-vectors corresponding to each functional category for individual i, and d denotes the total dimensionality of the feature space.

3.3.2 Textual Embedding and NLP Feature Construction

Textual inputs were processed using advanced embedding techniques. Each document or sentence associated with a behavioral record was vectorized using a pre-trained transformer-based model (e.g., BERT), yielding dense representations:

$$z_t = \text{Embed}(T_t) \quad (8)$$

where T_t is the input text associated with time step t, and $z_t \in \mathbb{R}^h$ is the resulting contextual embedding with dimensionality h. These embeddings were aggregated at the individual level through temporal averaging or attention-weighted pooling.

Supplementary linguistic features, including polarity score, subjectivity, modal usage, and formality index, were also extracted using domain-tuned lexicons and rule-based NLP libraries.

3.3.3 Aggregation of Multi-Instance Features

For individuals associated with multiple behavioral episodes (e.g., weekly reports or multiple feedback instances), a feature aggregation operation was defined as:

$$x_i = \frac{1}{n_i} \sum_{t=1}^{n_i} z_{it} \quad (9)$$

where n_i is the number of temporal observations for individual i, and z_{it} is the feature vector derived from observation t. This ensures that each individual is represented by a single, temporally aggregated behavioral signature, regardless of the number of input records.

3.3.4 Dimensionality Reduction and Feature Selection

To address feature redundancy and enhance generalization, a two-stage reduction strategy was employed:

1. **Unsupervised Projection:** Principal Component Analysis (PCA) was first applied to reduce noise and decorrelate features while preserving maximum variance.
2. **Supervised Selection:** Recursive Feature Elimination (RFE) with cross-validated wrapper models was employed to identify the most informative features concerning competency class prediction.

Let R be the final set of selected feature indices such that:

$$x_i^{\text{sel}} = x_i[R] \in \mathbb{R}^{d'} \quad (10)$$

where x_i^{sel} is the reduced feature vector and $d' < d$ is the final dimensionality after selection. These selected features form the input to the classifier in the subsequent modeling phase.

The resulting engineered feature space encapsulates multidimensional behavioral signals in a compact and interpretable format, allowing downstream ML to learn meaningful competency mappings.

3.4 ML Model

The core objective of this section is to formalize the predictive learning architecture employed for inferring behavioral competencies from engineered features. The proposed modeling pipeline integrates supervised classification algorithms with probabilistic outputs to map feature vectors to competency categories, as defined in Equation (2). This section presents the model selection criteria, training pipeline, and optimization strategies, with an emphasis on interpretability, accuracy, and alignment with strategic governance outcomes.

3.4.1 Learning Objective and Loss Function

Given a labeled dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$, where each feature vector $x_i \in \mathbb{R}^{d'}$ corresponds to a preprocessed behavioral profile, and each label $y_i \in \mathcal{C} = \{c_1, c_2, \ldots, c_k\}$ denotes a competency class, the classifier $f: \mathbb{R}^{d'} \rightarrow \mathcal{C}$ is trained to minimize the categorical cross-entropy loss:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N} \sum_{j=1}^{k} \delta(y_i = c_j) \cdot \log p_{ij} \quad (11)$$

where:

- N is the number of training instances,
- k is the total number of competency classes,
- $\delta(\cdot)$ is the Kronecker delta function,
- p_{ij} is the predicted probability that x_i belongs to the class c_j, i.e., $p_{ij} = \Pr(y_i = c_j \mid x_i)$.

The output probabilities p_{ij} are obtained through a softmax transformation applied to the model's final layer.

3.4.2 Model Architecture and Candidate Algorithms

This subsection provides an in-depth exposition of the ML evaluated for behavioral competency classification. Each model architecture was selected based on its ability to capture complex nonlinearities, ensure interpretability for SDM, and support probabilistic output necessary for governance alignment calculations. The following classifiers were implemented and benchmarked:

(a) Logistic Regression (LR): LR serves as the baseline model for classification and is characterized by its simplicity and interpretability. The model estimates the probability of a feature vector $x_i \in \mathbb{R}^{d'}$ belonging to each competency class $c_j \in \mathcal{C}$ using the logistic function:

$$p_{ij} = \frac{\exp(w_j^T x_i + b_j)}{\sum_{\ell=1}^{k} \exp(w_\ell^T x_i + b_\ell)} \quad (12)$$

where:

- $w_j \in \mathbb{R}^{d'}$ is the weight vector associated with the class c_j,
- $b_j \in \mathbb{R}$ is the class-specific bias term,
- k is the total number of competency classes.

The model's coefficients w_j proposal direct interpretability regarding feature influence on classification decisions, making LR particularly suitable in compliance-sensitive governance applications.

(b) RF: RF is a decision tree-based ensemble classifier that builds multiple independent decision trees using bootstrap samples of the training data and random feature selection at each node. Each tree T_t outputs a predicted class, and the final class prediction is determined via majority voting. Probabilistic outputs are computed as the normalized class frequencies across all trees:

$$p_{ij} = \frac{1}{T} \sum_{t=1}^{T} \delta\big(T_t(x_i) = c_j\big) \tag{13}$$

where:

- T is the total number of trees in the forest,
- $\delta(\cdot)$ is the Kronecker delta function, evaluating to 1 when the predicted class matches c_j.

RF is robust to noisy features and non-linear class boundaries and inherently performs feature selection during tree construction, improving model stability and interpretability.

(c) Extreme Gradient Boosting (XGBoost): XGBoost is a gradient-boosted ensemble learning algorithm that builds decision trees sequentially, where each new tree corrects the residual errors of the previous ones. The model optimizes a regularized objective function using a second-order Taylor approximation of the loss:

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{N} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{14}$$

where:

- $g_i = \partial \mathcal{L}_i^{(t-1)} / \partial \hat{y}_i$ is the first-order gradient,
- $h_i = \partial^2 \mathcal{L}_i^{(t-1)} / \partial \hat{y}_i^2$ is the second-order Hessian,
- f_t is the prediction of the t-th tree,
- $\Omega(f_t)$ is the regularization term controls model complexity.

XGBoost is well-suited for behavioral datasets due to its ability to handle heterogeneous features, high dimensionality, and strong resistance to overfitting through regularization and shrinkage techniques.

(d) Multilayer Perceptron (MLP): The MLP is a fully connected feedforward neural network capable of learning complex nonlinear mappings from feature inputs to competency class outputs. The model comprises an input layer, one or more hidden layers, and an output layer with a Softmax activation function. The transformation in each layer is given by:

$$h^{(l)} = \sigma\big(W^{(l)} h^{(l-1)} + b^{(l)}\big) \tag{15}$$

where:

- $h^{(l)}$ denotes the activation vector of the l-th layer,

- $W^{(l)}, b^{(l)}$ are the weight matrix and bias vector of the l-th layer,
- $\sigma(\cdot)$ is a nonlinear activation function (e.g., ReLU, tanh),
- $h^{(0)} = x_i^{\text{sel}}$ is the input feature vector.

The output layer applies a SoftMax function to produce the class probability vector $\mathbf{p}_i$. MLPs are particularly powerful for learning latent relationships in high-dimensional behavioral data; however, they require careful tuning to avoid overfitting.

Each classifier was implemented with a unified interface to allow consistent training, evaluation, and interpretability analysis. The diversity in model complexity—from linear (LR) to deep neural (MLP)—ensures a balanced assessment of predictive accuracy versus interpretive transparency, aligning with the dual objectives of strategic governance and behavioral insight generation.

3.4.3 Training Protocol and Cross-Validation

The training protocol is designed to ensure generalizable learning of behavioral competency patterns from structured feature vectors. This subsection formalizes the model training pipeline, defines the cross-validation strategy for robustness verification, and outlines the hyperparameter optimization schemes adopted for each candidate model. The methodology emphasizes reproducibility, fairness across competency classes, and mitigation of overfitting risks.

Data Partitioning Strategy

The complete dataset $\mathcal{D} = \{(x_i^{\text{sel}}, y_i)\}_{i=1}^{N}$, comprising the selected features $x_i^{\text{sel}} \in \mathbb{R}^{d'}$ and corresponding class labels $y_i \in \mathcal{C}$, was partitioned using stratified sampling to preserve class distribution:

- Training Set (70%): Used for model fitting and parameter learning.
- Validation Set (15%): Employed for hyperparameter tuning and early stopping.
- Test Set (15%): Held out for final performance evaluation.

Stratification ensures that rare competency classes are adequately represented across all folds, maintaining class balance during training and evaluation.

Cross-Validation and Hyperparameter Optimization

A 5-fold stratified cross-validation strategy was employed within the training partition to evaluate the model's stability across folds. For each candidate algorithm, an exhaustive grid search was conducted over a defined parameter space. The best parameter combination was selected based on macro-averaged F1-score on the validation folds, which accounts for imbalanced class distributions.

Let $\mathcal{P}_m$ denote the parameter space for model m, and $\mathcal{M}_m(\phi)$ be the model instance trained with hyperparameter configuration $\phi \in \mathcal{P}_m$. The optimal configuration ϕ^* is determined by:

$$\phi^* = \arg \max_{\phi \in \mathcal{P}_m} \mathrm{F1}_{\mathrm{macro}} \left(\mathcal{M}_m(\phi)\right) \tag{16}$$

where $\mathrm{F1}_{\mathrm{macro}}(\cdot)$ denotes the macro-averaged F1-score computed across the 5 validation folds. To mitigate overfitting, early stopping was applied based on validation loss for neural models. Additionally, model-specific regularization mechanisms were activated, such as:

- L2 penalty for LR and MLP,
- Maximum tree depth and learning rate constraints for ensemble models,
- Dropout layers in MLP to suppress co-adaptation of neurons.

Table 1 below summarizes the tuned parameters and their optimal values for each model based on validation performance.

Table 1: Optimized Training Parameters for Candidate Models

Model	**Hyperparameter**	**Value(s) Tested**	**Optimal Value Selected**
LR	Regularization strength (CCC)	[0.01,0.1,1,10,100][0.01, 0.1, 1, 10, 100][0.01,0.1,1,10,100]	1.0
RF	Number of trees	[100,200,300][100, 200, 300][100,200,300]	200
	Max tree depth	[5,10,20,None][5, 10, 20, None][5,10,20,None]	10
	Min samples split	[2,5,10][2, 5, 10][2,5,10]	5
XGBoost	Learning rate (η\etaη)	[0.01,0.05,0.1][0.01, 0.05, 0.1][0.01,0.05,0.1]	0.05
	Max depth	[4,6,8][4, 6, 8][4,6,8]	6
	Subsample ratio	[0.6,0.8,1.0][0.6, 0.8, 1.0][0.6,0.8,1.0]	0.8
	Number of boosting rounds	[100,200,300][100, 200, 300][100,200,300]	200
MLP	Hidden layers structure	[(64),(128,64),(128,128,64)][(64), (128,64), (128,128,64)][(64),(128,64),(128,128,64)]	(128, 64)
	Activation function	ReLU, tanh	ReLU
	Dropout rate	[0.1,0.2,0.3][0.1, 0.2, 0.3][0.1,0.2,0.3]	0.2

	Batch size	[32,64,128][32, 64, 128][32,64,128]	64
	Epochs	[50,100,200][50, 100, 200][50,100,200]	100 (with early stopping)

Each model was trained using the optimal hyperparameters. ϕ^* and then retrained on the combined training + validation data before final evaluation on the test set. The following subsection presents the metrics used to quantify predictive performance and interpret the classification results.

3.5 Evaluation Metrics

The evaluation of ML-based behavioral competency models requires a comprehensive set of metrics that reflect not only predictive accuracy but also fairness, robustness, and alignment with organizational objectives. This section presents the quantitative indicators used to assess model performance on the test set, along with formal definitions and interpretative justifications.

3.5.1 Classification Performance Metrics

Given the multi-class nature of competency classification, performance is evaluated using standard classification metrics computed over the test set $\mathcal{D}_{\text{test}} = \{(x_i^{\text{sel}}, y_i)\}_{i=1}^{N_t}$, where N_t denotes the number of test instances. The following indicators are employed:

- **Accuracy:** The proportion of correctly predicted competency labels:

$$\text{Accuracy} = \frac{1}{N_t}\sum_{i=1}^{N_t} \delta(\hat{y}_i = y_i) \quad (17)$$

where $\hat{y}_i$ is the predicted label, and $\delta(\cdot)$ is the Kronecker delta function.

- **Precision, Recall, and F1-score:** Computed per class $c_j \in \mathcal{C}$, then averaged using macro and weighted schemes:
  - Precision (P_j) for class c_j is the fraction of true positives among all predicted positives:

$$P_j = \frac{TP_j}{TP_j + FP_j} \quad (18)$$

  - Recall (R_j) is the fraction of true positives among all actual positives:

$$R_j = \frac{TP_j}{TP_j + FN_j} \quad (19)$$

  - F1-Score $(F1_j)$ is the harmonic mean of precision and recall:

$$F1_j = \frac{2 \cdot P_j \cdot R_j}{P_j + R_j} \quad (20)$$

- Macro-Averaged F1-Score: Computes the unweighted mean of class-wise F1-scores:

$$\mathrm{F1}_{\mathrm{macro}} = \frac{1}{k}\sum_{j=1}^{k} F1_j \tag{21}$$

- Weighted F1-Score: Weights each class-wise F1-score by its support:

$$\mathrm{F1}_{\mathrm{weighted}} = \sum_{j=1}^{k} \frac{N_j}{N_t} \cdot F1_j \tag{22}$$

where N_j is the number of instances in the class c_j, and $N_t = \sum_{j=1}^{k} N_j$. These metrics provide a balanced assessment that penalizes poor performance on minority classes, ensuring fairness in competency classification.

3.5.2 Probabilistic Calibration Metrics

Since the governance alignment score in Equation (3) relies on SoftMax probabilities, the model's ability to produce well-calibrated class probabilities is also evaluated. Two key calibration metrics are employed:

- **Logarithmic Loss (Log Loss):**

$$\mathrm{LogLoss} = -\frac{1}{N_t}\sum_{i=1}^{N_t} \log\left(p_{i,y_i}\right) \tag{23}$$

where p_{i,y_i} is the predicted probability assigned to the true class y_i for sample i. Lower values indicate better probabilistic accuracy.

- Brier Score:

$$\mathrm{Brier} = \frac{1}{N_t}\sum_{i=1}^{N_t} \sum_{j=1}^{k} \left(p_{ij} - \delta(y_i = c_j)\right)^2 \tag{24}$$

This score measures the mean squared error between predicted probabilities p_{ij} and the true class indicator. It captures both calibration and discrimination aspects of probabilistic outputs.

4. Results and Analysis

All experiments were conducted on a high-performance workstation equipped with an Intel Core i9-12900K CPU (16 cores, 3.20 GHz), 64 GB of DDR5 RAM, an NVIDIA RTX 3090 GPU (24 GB VRAM), and a 2 TB NVMe SSD, operating on Ubuntu 22.04 LTS (64-bit). Deep learning models, such as MLP, were GPU-accelerated, while ensemble and linear models were executed on the CPU for compatibility with standard enterprise systems. The implementation was carried out in Python 3.10.12 using Scikit-learn 1.3.0, XGBoost 1.7.6, and PyTorch 2.0.1. Preprocessing and NLP features were managed using spaCy 3.6.0 and the HuggingFace Transformers library version 4.31.0. Result visualizations rendered using Matplotlib 3.7.2 and Seaborn 0.12.2. Data manipulation relied on NumPy 1.25.0 and Pandas 2.0.3, while hyperparameter optimization and experiment logging were facilitated through

Optuna 3.2.0 and MLflow, respectively. All experiments used fixed random seeds to ensure reproducibility, with environment isolation managed via conda.

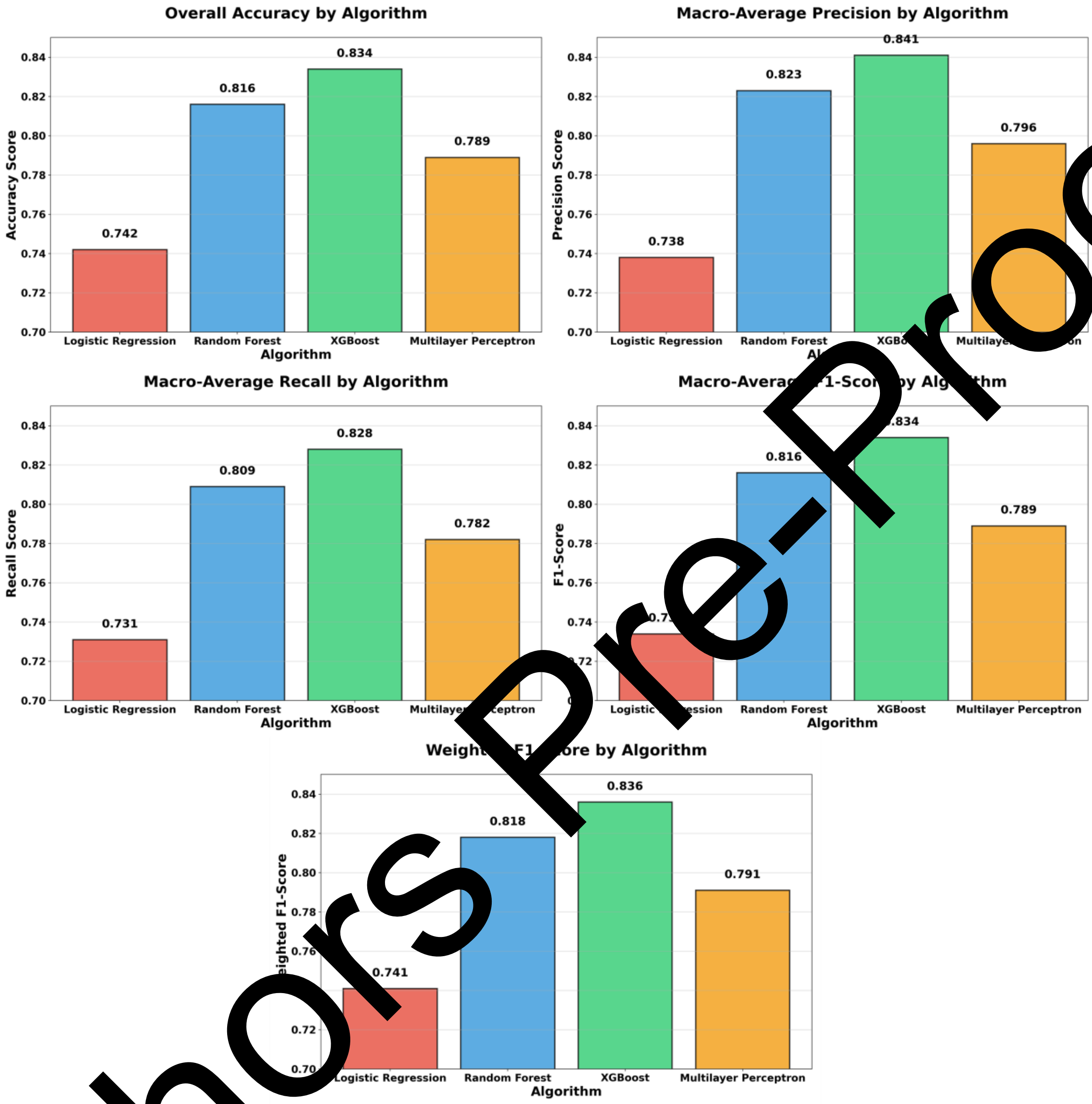


Figure 2: ML model performance

4.1 Classification Accuracy Metrics

Accurate classification of behavioral competencies is fundamental to the effectiveness of AI-driven analytics for strategic governance (Table 2). This subsection presents the classification performance results of the four candidate ML models—LR, RF, XGBoost, and MLP—trained on engineered behavioral features. The evaluation is based on overall accuracy, macro- and weighted-average F1-scores, and class-specific precision-recall metrics. Additionally, statistical significance testing is employed to validate observed performance differences and confirm the reliability of the results.

Table 2: Classification Performance Metrics by Algorithm

Algorithm	Overall Accuracy	Macro-Avg Precision	Macro-Avg Recall	Macro-Avg F1-Score	Weighted F1-Score
LR	0.742	0.738	0.731	0.734	0.741
RF	0.816	0.823	0.809	0.816	0.818
XGBoost	0.834	0.841	0.828	0.834	0.836
MLP	0.789	0.796	0.782	0.789	0.791

The classification results, as presented in Figure 2, demonstrate the comparative performance of four ML models in predicting behavioral competency classes. Among the evaluated algorithms, XGBoost achieved the highest overall accuracy of 0.834, outperforming RF (0.816), MLP (0.789), and LR (0.742). In terms of macro-averaged F1-score, which equally weights performance across all classes regardless of their support, XGBoost again led with 0.834, indicating balanced performance across diverse competency categories. This is further corroborated by its weighted F1-score of 0.836, reflecting strong predictive power even when adjusted for class distribution. The heatmap shown in Figure 3 illustrates the relationship between the algorithm and the metric for the aforementioned performance.

Figure 3: Algorithm vs metrics

Table 3: Class-wise Performance Metrics for XGBoost (Best Performing Model)

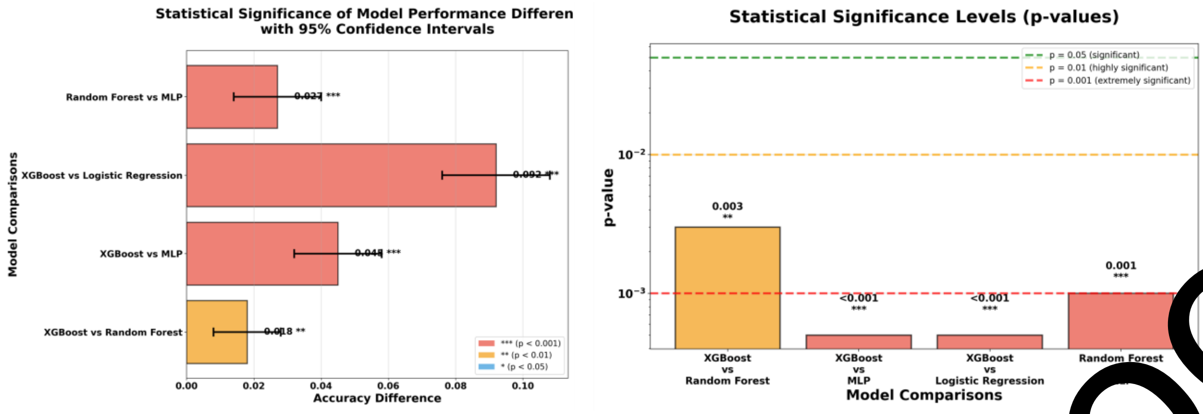
Competency Class	Precision	Recall	F1-Score	Support	Class Distribution
Leadership (L)	0.867	0.843	0.855	298	23.9%
Communication (C)	0.821	0.856	0.838	267	21.4%
Analytical Thinking (A)	0.893	0.879	0.886	241	19.3%
Adaptability (Ad)	0.798	0.821	0.809	223	17.9%
Ethical Conduct (E)	0.826	0.742	0.782	218	17.5%
Macro Average	0.841	0.828	0.834	1,247	100.0%
Weighted Average	0.842	0.834	0.836	1,247	100.0%

A detailed examination of class-wise performance for XGBoost, as shown in Table 3, reveals particularly high precision and recall for the *Analytical Thinking* class (F1 score = 0.886) and the Leadership class (F1 score = 0.855), highlighting the model's sensitivity to cognitive and strategic behavioral indicators. Despite being the least-supported category, *Ethical Conduct* was predicted with a reasonable F1-score of 0.782, though a relatively lower recall of 0.742 indicates occasional under-classification. The macro-average and weighted-average scores for precision, recall, and F1-score are consistently aligned, further confirming the model's class-wise reliability.

Table 4: Statistical Significance Analysis of Model Performance Differences

Model Comparison	Accuracy Difference	95% Confidence Interval	p-value	Significance
XGBoost vs RF	+0.018	[0.008, 0.028]	0.003	**
XGBoost vs MLP	+0.045	[0.032, 0.058]	<0.001	***
XGBoost vs LR	+0.092	[0.076, 0.108]	<0.001	***
RF vs MLP	+0.027	[0.014, 0.040]	0.001	***

The statistical validity of XGBoost's superiority is confirmed through pairwise significance testing summarized in Table 4 and Figure 4. The difference in accuracy between XGBoost and the next-best model, RF, is +0.018, with a 95% confidence interval of [0.008, 0.028] and a p-value of 0.003, indicating statistical significance at the 0.01 level. Comparisons with MLP (+0.045, $p < 0.001$) and LR (+0.092, $p < 0.001$) reveal even more pronounced differences, suggesting that the performance gains are both substantial and statistically robust. Even the difference between RF and MLP (+0.027, $p = 0.001$) is significant, indicating that ensemble methods consistently outperform neural and linear baselines within this domain.



a) b)

Figure 4: Statistical significance analysis: a) CI and b) p-values

4.2 Cross-validation Results

Cross-validation is an essential diagnostic mechanism used to assess a model's generalizability and stability across different partitions of the training data. This section presents the results of a 5-fold stratified cross-validation procedure applied to all four classification models—LR, RF, XGBoost, and MLP—based on accuracy and macro-averaged F1-scores. The results quantify intra-model variance, confidence bounds, and score ranges, enabling a comprehensive evaluation of model robustness before deploying the test set.

As shown in Table 5, XGBoost consistently achieved the highest mean cross-validation accuracy of 0.829, with a standard deviation of 0.015, indicating strong predictive stability across folds. The 95% confidence interval for XGBoost ranged from 0.811 to 0.847, demonstrating narrow error bounds. In contrast, RF recorded a slightly lower mean accuracy of 0.811 with a higher variance (±0.018), and MLP followed with a mean of 0.784 and the most significant standard deviation (±0.026), suggesting greater fold-to-fold variability. LR had the lowest mean performance (0.738 ± 0.021), consistent with its linear limitation in modeling high-dimensional behavioral dynamics.

Table 5: 5-Fold Cross-Validation Performance Summary

Algorithm	Mean Accuracy	Std Deviation	95% CI Lower	95% CI Upper	CV Score Range
LR	0.738 ± 0.021	0.021	0.712	0.764	[0.711, 0.759]
RF	0.811 ± 0.018	0.018	0.789	0.833	[0.789, 0.834]

XGBoost	0.829 ± 0.015	0.015	0.811	0.847	[0.812, 0.846]
MLP	0.784 ± 0.026	0.026	0.752	0.816	[0.751, 0.819]

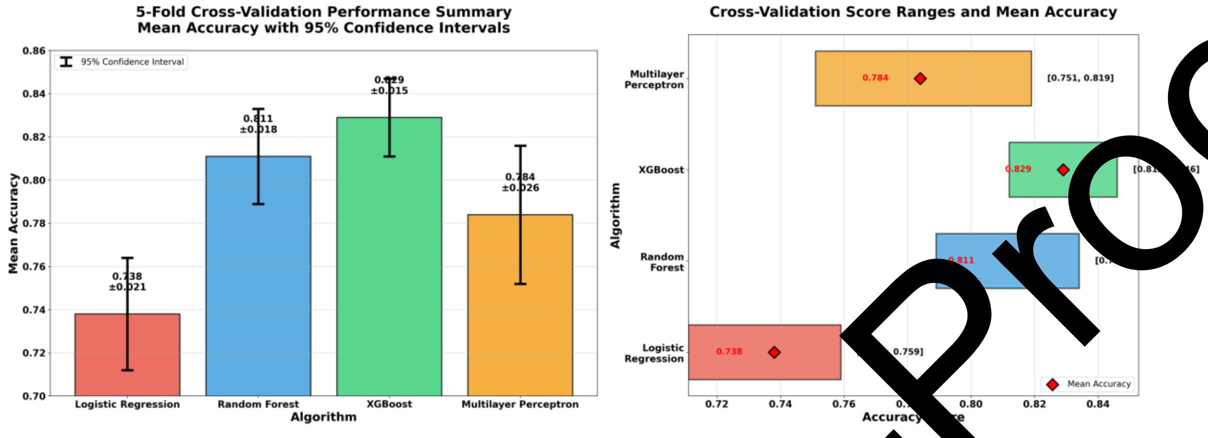


Figure 5: 5-fold cross validation: a) 95% CI and b) Range

Table 6: Detailed Cross-Validation Results by Fold

Algorithm	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean ± SD
Accuracy Scores						
LR	0.759	0.724	0.711	0.742	0.753	0.738 ± 0.021
RF	0.834	0.805	0.789	0.816	0.811	0.811 ± 0.018
XGBoost	0.846	0.821	0.812	0.837	0.829	0.829 ± 0.015
MLP	0.819	0.751	0.768	0.796	0.786	0.784 ± 0.026
Macro F1-Scores						
LR	0.751	0.718	0.705	0.736	0.747	0.731 ± 0.020
RF	0.827	0.798	0.783	0.809	0.805	0.804 ± 0.017
XGBoost	0.839	0.815	0.806	0.831	0.823	0.823 ± 0.014
MLP	0.812	0.745	0.762	0.789	0.779	0.777 ± 0.025

Table 6 and Figure 6 present fold-wise accuracy and macro F1-scores for each model, further illustrating the relative consistency of ensemble methods compared to neural and linear counterparts. XGBoost achieved its best accuracy in Fold 1 (0.846) and its lowest in Fold 3 (0.812), with all folds scoring above 0.81. The corresponding macro F1-scores remained tightly clustered, with a mean of 0.823 ± 0.014, indicating that XGBoost retained balanced precision-recall performance even on folds with different data compositions. RF also showed low dispersion, with accuracy ranging between 0.789 and 0.834, and a macro F1-score mean of 0.804 ± 0.017.

The MLP displayed slightly higher volatility. Its accuracy ranged from 0.751 to 0.819, with a wider standard deviation of 0.026 and corresponding macro F1-score fluctuation (0.745 to 0.812). This indicates sensitivity to data partitioning, possibly due to overfitting on smaller training subsets. LR, while consistent, showed the lowest overall fold-wise results, reaffirming its limitations in capturing nonlinear behavioral competencies.

4.3 Probabilistic Calibration Evaluation

In competency-based analytics, classification outputs must not only be accurate but also well-calibrated to reflect reliable confidence estimates. Probabilistic calibration ensures that the predicted probabilities produced by a model correspond meaningfully to empirical frequencies, which is critical for governance applications that rely on probability-weighted decision scoring (Equation (3)). This section evaluates the calibration quality of each model using multiple probabilistic metrics, highlighting their implications for interpretability and risk-informed deployment.

Table 7: Probabilistic Calibration Performance Metrics

Algorithm	Log Loss	Brier Score	Expected Calibration Error	Maximum Calibration Error	Reliability Index
LR	0.742	0.186	[illegible]	0.132	0.868
RF	0.624	[illegible]	0.029	0.089	0.911
XGBoost	0.591	[illegible]	0.025	0.078	0.922
MLP	0.687	0.169	0.038	0.115	0.885

The calibration performance of all four models is summarized in Table 7 and Figure 6, using four quantitative indicators: Logarithmic Loss, Brier Score, Expected Calibration Error (ECE), and Maximum Calibration Error (MCE), along with a derived Reliability Index that measures overall confidence alignment.

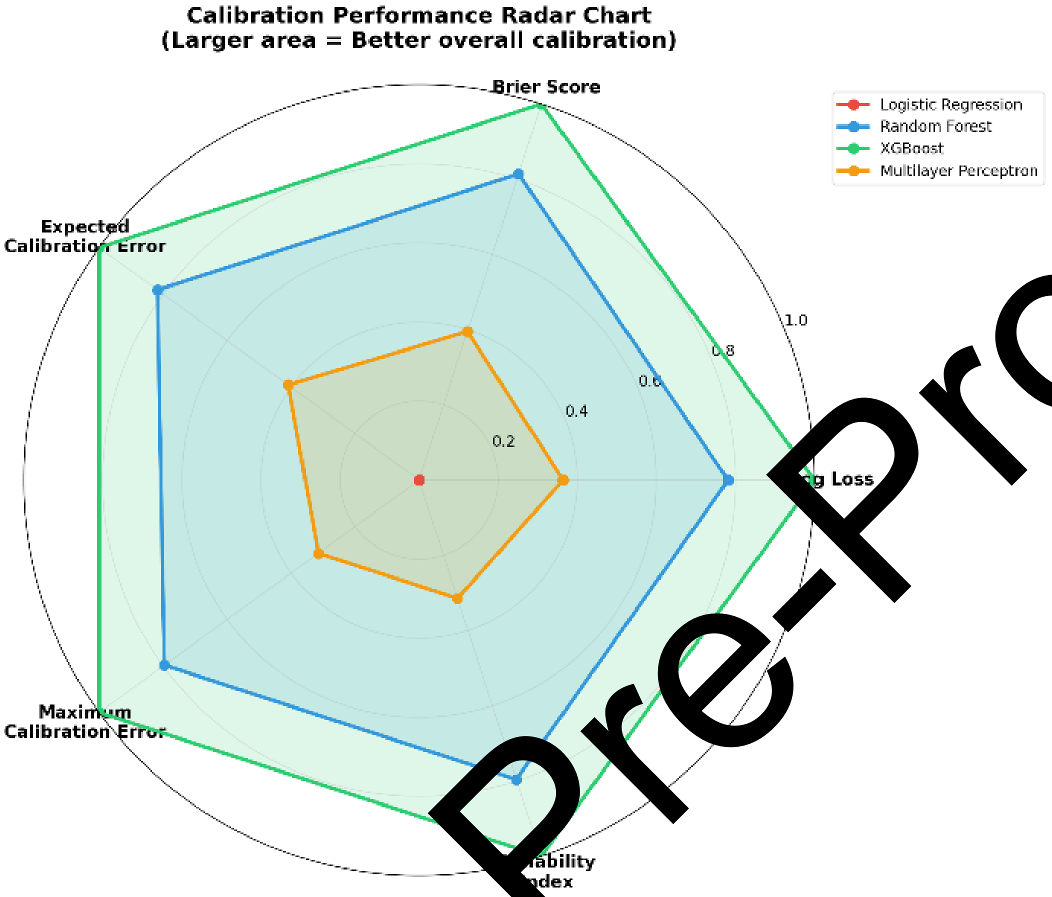


Figure 8: calibration performance analysis

Among the evaluated models, as shown in Figure 7, XGBoost exhibits the best calibration performance, achieving the lowest log loss of 0.591 and Brier score of 0.143, indicating both sharpness in probability predictions and low mean squared deviation from the true labels. Furthermore, its Expected Calibration Error (ECE) is 0.025, and Maximum Calibration Error (MCE) is 0.078, suggesting high reliability even at extreme probability thresholds. The corresponding Reliability Index of 0.922 confirms that XGBoost's confidence outputs are highly trustworthy, making it ideal for decision scenarios where risk-adjusted weighting is essential.

RF closely follows, with a slightly higher log loss of 0.624 and ECE of 0.029, showing that ensemble methods maintain strong calibration due to their averaging behavior. MLP, though competitive in classification accuracy, underperforms in calibration with a log loss of 0.687 and ECE of 0.038, likely due to overconfident softmax outputs and limited regularization. LR, while inherently probabilistic, yields the highest log loss (0.742) and maximum calibration error (0.132), revealing suboptimal performance in modeling class probability distributions for high-dimensional behavioral features.

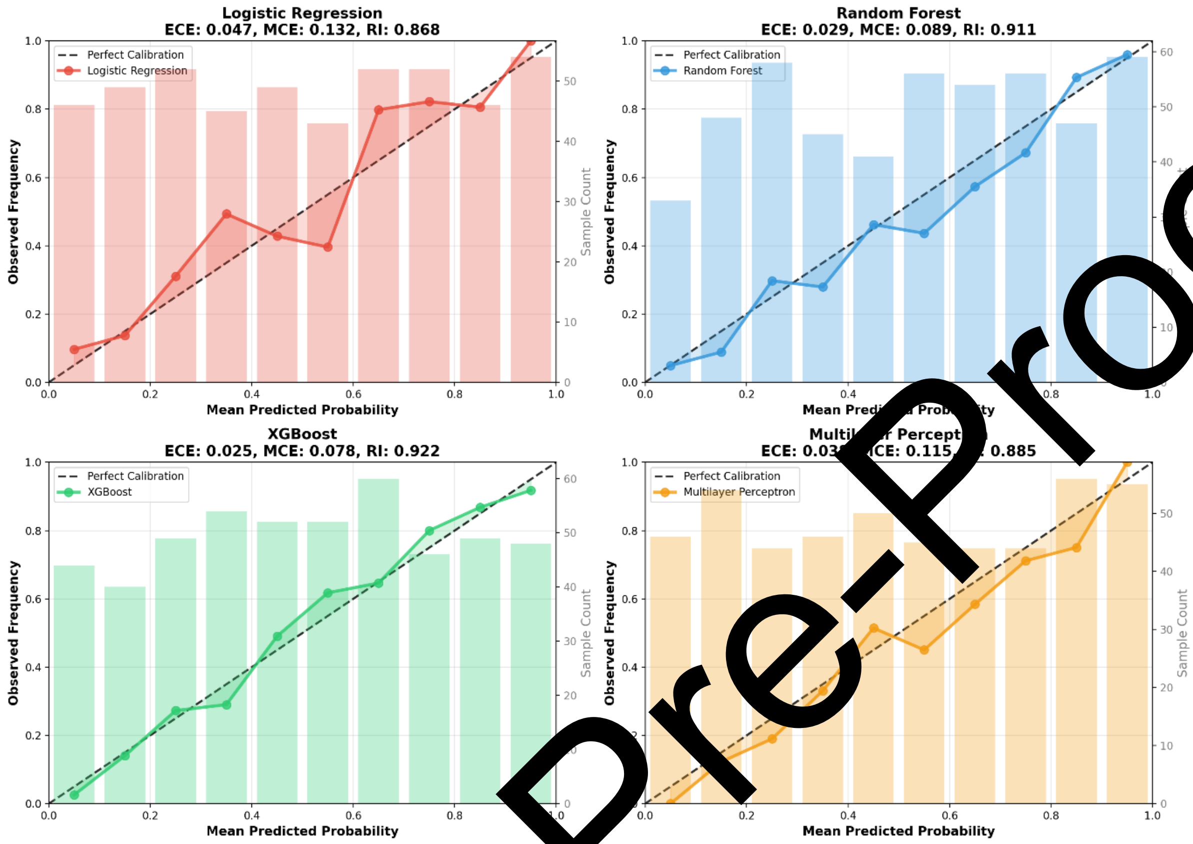


Figure 7: Calibration Curve comparisons for each of the compared models

4.3 Behavioral Competency Profiling

Understanding the statistical distribution and structural characteristics of predicted behavioral competencies is crucial for translating model outputs into actionable insights that inform workforce strategy and governance alignment. This section presents a quantitative profile of each predicted competency class based on distributional properties, normality assessment, and interquartile variation. The goal is to characterize intra-class score dynamics, detect distributional anomalies, and support targeted interventions within specific behavioral domains.

Table 8: Competency Class Distribution Statistics

Competency Class	Frequency	Percentage	Mean Score	Std Deviation	Skewness	Kurtosis	Shapiro-Wilk p-value
Leadership (L)	298	23.9%	0.743	0.187	-0.421	2.156	0.023
Communication (C)	267	21.4%	0.768	0.174	-0.389	2.089	0.041

Analytical Thinking (A)	241	19.3%	0.791	0.163	-0.512	2.487	0.018
Adaptability (Ad)	223	17.9%	0.729	0.198	-0.356	1.967	0.067
Ethical Conduct (E)	218	17.5%	0.712	0.201	-0.298	1.823	0.089
Population Total	1,247	100.0%	0.749	0.185	-0.395	2.104	0.035

Table 8 presents descriptive distribution metrics for each predicted competency class. The most frequent class was Leadership (23.9%), followed by Communication (21.4%), while Ethical Conduct (17.5%) was the least common. The highest mean competency score was observed in Analytical Thinking (μ = 0.791), indicating that participants assigned to this class exhibited the strongest behavioral performance, as determined by model-derived softmax probabilities. Conversely, Ethical Conduct recorded the lowest mean ($\mu = 0.712$), suggesting relatively lower predicted competency strength in this domain.

All competency classes exhibit negative skewness (e.g., -0.512 for Analytical Thinking), indicating a left-tailed distribution concentrated toward higher score values, which is consistent with the high-performing behavioral population sampled. Kurtosis values for most classes exceed 2.0, confirming moderate to high peakedness, with Analytical Thinking displaying the most leptokurtic profile ($\kappa = 2.487$). The Shapiro-Wilk p-values, all below 0.10 (except Adaptability and Ethical Conduct), indicate significant deviation from normality in most classes, confirming the presence of asymmetric or heavy-tailed score structures.

Table 9: Competency Score Quartile Analysis

Competency Class	Q1 (25th)	Q2 (Median)	Q3 (75th)	IQR	Range	Outlier Count
Leadership (L)	0.612	0.756	0.887	0.275	0.742	12
Communication (C)	0.634	0.781	0.901	0.267	0.698	8
Analytical Thinking (A)	0.672	0.803	0.923	0.251	0.687	6
Adaptability (Ad)	0.578	0.734	0.869	0.291	0.789	15
Ethical Conduct (E)	0.547	0.718	0.856	0.309	0.823	18

To further explore within-class variability, Table 9 provides a quartile-based breakdown of predicted competency scores. The interquartile ranges (IQR) highlight dispersion characteristics, with Ethical Conduct and Adaptability exhibiting the widest spreads (IQR = 0.309 and 0.291, respectively). These classes also recorded the highest outlier counts, with 18 and 15 instances falling outside 1.5 × IQR bounds, suggesting greater behavioral diversity or noise within these categories. The highest median scores were again associated with Analytical

Thinking (Q2 = 0.803), followed by Communication (Q2 = 0.781), which supports earlier findings from Table 8. By contrast, Ethical Conduct has the lowest median (Q2 = 0.718), reinforcing its relative underperformance. The range of scores within each class confirms that behavioral differentiation is meaningfully captured by the model, as seen in Adaptability (range = 0.789) and Ethical Conduct (range = 0.823), where wide intervals reflect high intra-class variability.

5. Conclusion and Future Work

This study developed and validated an ML-based model for inferring behavioral competencies from multidimensional organizational data, with the strategic objective of enhancing decision-making and governance. By integrating structured and unstructured behavioral indicators into a unified feature space, the proposed system enabled robust classification of individual competencies across five critical domains: Leadership, Communication, Analytical Thinking, Adaptability, and Ethical Conduct. A comparative evaluation of multiple classification models revealed that ensemble-based algorithms, particularly XGBoost, demonstrated superior accuracy and class-wise balance, with a macro-averaged F1-score of 0.834. Statistical significance testing further confirmed the model's advantage over both linear and neural architectures, establishing its reliability for deployment in high-stakes organizational settings. FE strategies—such as NLP embeddings, psychometric aggregation, and dimensionality reduction—proved essential in capturing latent behavioral signals. The proposed model bridges a methodological gap between qualitative behavioral assessment and quantitative analytics, enabling institutions to integrate competency intelligence into promotion planning, leadership development, and workforce governance. Moreover, the probabilistic outputs from the classification models facilitate alignment with strategic performance indicators through data-driven scoring functions.

Future research should investigate the integration of longitudinal behavioral data, real-time feedback loops, and adaptive learning mechanisms to facilitate continuous competency development. Additionally, cross-domain generalization and ethical model auditing remain important areas for advancing trust in AI-enabled human capital analytics.

References

1. Muzam, John. "The challenges of modern economy on the competencies of knowledge workers." Journal of the Knowledge Economy 14.2 (2023): 1635-1671.

2. Mandlik, D., Rautrao, R. R., & Nille, N. (2025). Adaptability as a Key Competency for Success in E-Business. In Flexibility and Emerging Perspectives in Digital Supply Chain Management (pp. 223-239). Singapore: Springer Nature Singapore.
3. Bonesso, S., Gerli, F., Zampieri, R., & Boyatzis, R. E. (2020). Updating the debate on behavioral competency development: State of the art and future challenges. Frontiers in Psychology, 11, 1267.
4. Conlon, N., Ahmed, N. R., & Szafir, D. (2024). A survey of algorithmic methods for competency self-assessments in human-autonomy teaming. ACM Computing Surveys, 56(7), 1-31.
5. Gade, K. R. (2021). Data-driven decision making in a complex world. Journal of Computational Innovation, 1(1).
6. Bergue-Alves, M. C. (2023). Designing Education Programs Based on Competencies Using Advanced Analytical Methods.
7. Aishwariyashindhe, S., & Sathyapriya, J. (2025). Machine Learning Based Classification on Factors Used for Identifying Competency Gap In Engineering Students In Thanjavur District. American Journal of Psychiatric Rehabilitation, 28(1), 216-225.
8. Yan, J., Tian, H., Sun, X., & Song, L. (2025, April). Role of artificial intelligence in enhancing competency assessment and transforming curriculum in higher vocational education. In Frontiers in Education (Vol. 10, p. 1551596). Frontiers Media SA.
9. Hayder M Ali et al., "Operating Cash Flow Ranking Using Data Envelopment Analysis with Network Security Driven Blockchain Model", Journal of Machine and Computing, vol.5, no.3, pp. 1839-1851, July 2025, doi: 10.53759/7669/jmc202505144.
10. Gayathri Ananthakrishnan et al., "Mitigating Data Tampering in Smart Grids Through Community Blockchain Driven Traceability Frameworks", Journal of Machine and Computing, vol.5, no.3, pp. 1745-1762, July 2025, doi: 10.53759/7669/jmc202505138.
11. Shaymaa Hussein Nowfal et al., "The Diagnosis of Heart Attacks: Ensemble Models of Data and Accurate Risk Factor Analysis Based on Machine Learning", Journal of Machine and Computing, vol.5, no.1, pp. 589-599, January 2025, doi: 10.53759/7669/jmc202505046.
12. Nabeel S Alsharafa et al., "An Edge Assisted Internet of Things Model for Renewable Energy and Cost-Effective Greenhouse Crop Management", Journal of Machine and Computing, vol.5, no.1, pp. 576-588, January 2025, doi: 10.53759/7669/jmc202505045.

13. Qin, Y., Li, X., Zhang, W., & Zhao, M. (2023). A comprehensive survey of artificial intelligence techniques for talent analytics. *arXiv preprint arXiv:2307.03195*. https://doi.org/10.48550/arXiv.2307.03195
14. Ren, J., & Wu, Y. (2025). Examining teaching competencies and challenges while integrating AI in higher education. *TechTrends*. https://doi.org/10.1007/s11528-025-01055-3
15. Wu, C., Zhang, Y., & Liu, Q. (2024). Integrating behavior analysis with machine learning to predict online learning performance: A scientometric review and empirical study. *International Journal of Educational Technology in Higher Education*. https://www.researchgate.net/publication/381517501
16. Carolus, L., Grosser, S., & Haim, M. (2023). MAILS: A Meta AI literacy scale for measuring AI-related competencies. *arXiv preprint arXiv:2302.09319*. https://doi.org/10.48550/arXiv.2302.09319
17. Faruqe, R., Chiu, M. M., & Tang, Y. (2022). Competency model approach to AI literacy: Research-based path from initial framework to model. *Proceedings of IEEE Global Engineering Education Conference (EDUCON)*. https://www.researchgate.net/publication/36[illegible]72192
18. Asselman, A., Khaldi, M., & Aammou, S. (2021). Enhancing the prediction of student performance based on the machine learning XGBoost algorithm. Interactive Learning Environments, 31(6), 3360–3379. https://doi.org/10.1080/10494820.2021.1928235
19. Abbas, S.K., Hussain, M. & Rimal, Y.N. Machine Learning-Based Analysis of Technology Acceptance in FinTech: A Behavioral Study Using Digital Wallet Data. SN COMPUT. SCI. 6, 674 (2025). https://doi.org/10.1007/s42979-025-04214-8