

An Interpretable Random Forest with SHAP Explanations for Multiclass Skill Level Classification Model in Malaysia's Labour Market

Rabi'atul'adawiah Shabli^{1,2}, Ahmad Zia Ul-Saufie^{1*}, Arief Gusnanto³, and Nurain Ibrahim¹

¹Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

²Department of Statistics Malaysia, Ministry of Economy, Federal Government Administrative Centre, 62514 Putrajaya, Wilayah Persekutuan Putrajaya, Malaysia

³Department of Statistics, School of Mathematics, Faculty of Engineering and Physical Sciences, University of Leeds, LS2 9JT Leeds, United Kingdom

ABSTRACT

Skill level classification is essential for effective human capital interventions, wage policy formulation and long-term workforce development. However, the black-box nature of machine learning algorithms limits the transparency and interpretability required for policymaking. This study aims to develop an accurate and interpretable multiclass skill level classification model for the Malaysian labour market. The model was constructed using secondary data from the Department of Statistics Malaysia's 2023 Salaries and Wages Survey, covering 120,518 Malaysian workers. Data preprocessing involved cleaning, transformation and addressing class imbalance using the Synthetic Minority Oversampling Technique (SMOTE). Using the top nineteen XGBoost feature importance scores, a Random Forest model was selected due to its robust ensemble mechanism and trained using these features. The proposed model exhibits strong predictive performance across all metrics,

achieving an accuracy of 0.8754, sensitivity of 0.8485, specificity of 0.9384 and F1-score of 0.8378. The interpretability of the most influential predictors was investigated using SHapley Additive exPlanations (SHAP) at both global and local levels, revealing the importance of salaries and wages activity, highest certificate and education level. These features provide deeper insights into understanding skill level classification across skilled, semi-skilled and low-skilled categories. The findings emphasise the significance of combining high predictive

ARTICLE INFO

Article history:

Received: 18 October 2025

Accepted: 04 March 2026

Published: 30 April 2026

DOI: <https://doi.org/10.47836/pjst.34.2.25>

E-mail addresses:

adawiah.dosm@gmail.com (Rabi'atul'adawiah Shabli)

ahmadzia101@uitm.edu.my (Ahmad Zia Ul-Saufie)

a.gusnanto@leeds.ac.uk (Arief Gusnanto)

nurainibrahim@uitm.edu.my (Nurain Ibrahim)

* Corresponding author

accuracy with transparent feature interpretability in labour market analysis. This study represents one of the initial labour market studies in Malaysia integrating Random Forest with SHAP. The proposed interpretable skill level framework can be utilised to facilitate evidence-based policymaking, targeted human capital interventions and workforce development planning in Malaysia.

Keywords: Classification, explainable artificial intelligence, feature selection, labour market, random forest, SHAP, skill level, XGBoost feature importance

INTRODUCTION

Machine learning algorithms are extensively applied to complex prediction tasks, yet they are often criticised for functioning as “black boxes”, where linkages between input features and target variables remain opaque (Neubauer et al., 2025). This lack of transparency posed a challenge among policymakers and decision-makers who required not only accurate predictions but also clear interpretability to cultivate trust and accountability. As a result, researchers have aimed to enhance machine learning algorithms by combining high predictive accuracy and techniques that can clarify the model behaviour in a way that is easily accessible and actionable (Neubauer et al., 2025; Sani et al., 2025).

Explainable Artificial Intelligence (XAI) has become an effective solution to address this issue, as this technique allows the operational mechanisms of black-box models to be made more transparent (Dhungana, 2025). Implementation of XAI will foster user confidence, model refinement and strengthen the governance framework. A considerable amount of literature has been published on the utilisation of interpretability techniques for classification purposes, such as SHapley Additive exPlanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME), Explain Like I'm Five (ELI5), Quantum Lattice (QLattice), and Anchor (Arravalli et al., 2025).

Nowadays, SHAP has been successfully integrated with machine learning algorithms in various fields. SHAP has been employed to clarify input predictors of therapy dropout (Visibelli et al., 2025), chronic kidney disease (Elshewey et al., 2025), and breast cancer symptoms (Arravalli et al., 2025). In medicine, it enables clinicians to discover non-linear associations among the features and facilitate precision medicine. Likewise, Sani et al. (2025) adopted the Random Forest and SHAP to clarify sociodemographic characteristics that affected the fertility preferences of Somalian women. SHAP has been used in the classification of industrial safety, including ExAIRFC-GSDC, to identify input predictors of gas leakage accidents with an accuracy of more than 98 per cent (Lalithadevi & Krishnaveni, 2025). Taken together, this application suggests that integrating SHAP can strengthen predictive accuracy and interpretability, with the intention that predictions are not only strong but also explainable.

Recent labour market studies that apply XAI methods have indicated an increase in the ability to overcome the limitations of traditional statistical analysis. This can be seen in other studies that have applied SHAP interpretation on the job losses among the elderly (Celbiş et al., 2023), unemployment rates in Europe (Celbiş, 2023), occupational injuries in South Korea (Lee et al., 2023), gig workers' employment (Liu & Wang, 2025), employee attrition (Manafi Varkiani et al., 2025) and online vacancies (Dossche et al., 2025). Together, these studies outline that tree-based algorithms, specifically Random Forest and XGBoost were found to be the most appropriate techniques due to their ability to predict, while SHAP enables revealing of the input predictors, including age, education, activity and employment status.

Taken together, prior research underscores the growing relevance of SHAP-enhanced machine learning in fields where both predictive performance and interpretability are essential. Until recently, amidst advances in medicine, industrial safety, workplace injuries and the labour market, there has been little research concentrated on the classification of skill level. Accurate and interpretable categorisation of workers into skilled, semi-skilled and low-skilled is vital for shaping wage policies, human capital interventions and long-term workforce planning. This study, therefore, set out to address the gap by combining the Random Forest with SHAP analysis to enhance skill level prediction in Malaysia. By combining strong predictive performance with feature-level transparency, the study aimed to provide actionable insights for policymakers to design evidence-based and targeted labour market initiatives.

Consequently, this study provides several key contributions. From the methodological perspective, this study incorporates an ensemble Random Forest model with SHAP explanations, enabling an accurate prediction and interpretability for multiclass skill level classification. Additionally, this study leverages national representative data of the 2023 Salaries and Wages Survey, consisting of 120,518 Malaysian workers, to identify the most influential skilled, semi-skilled and low-skilled input predictors across demographic, socioeconomic and employment-related features. Lastly, in terms of policy implications, this proposed skill level framework offers transparent and empirical information which may support human capital strategies pertaining to labour, employment and social protection for a dynamic and competitive Malaysian labour market.

METHODOLOGY

This study utilised the cross-sectional 2023 *Salaries and Wages Survey* dataset, a household survey approach of Labour Force/ Salaries & Wages/ Migration Survey from the Department of Statistics Malaysia (Department of Statistics Malaysia, 2024a). The dataset consisted of 120,518 workers with nineteen features and could be segmented into three components encompassing identification particulars, labour force particulars and salaries

and wages (Department of Statistics Malaysia, 2024a). It adopted a stratified two-stage sampling design to cater to the national representatives throughout every state across the nation (Department of Statistics Malaysia, 2024b). The research workflow was illustrated in Figure 1, which describes a sequence of processes encompassing data preprocessing, feature selection, model training and model interpretation.

This study is based on a cross-sectional dataset that captures the workforce in a single year of 2023. Given that the dataset is cross-sectional, the analysis assumes that the features reflect stable and underlying patterns in the labour market. The assumption of limited short-term mobility is supported by the Thirteenth Malaysia Plan 2026-2030, which indicates that social mobility indicators remain at a low level and that young workers continue to experience constraints in reaching the skills required for the job market (Ministry of Economy, 2025). This classification model serves as a structural benchmark that identifies fundamental requirements, such as education level and the highest certificate, to define skill level classification in the Malaysian labour market at a particular time.

Skill Level Categories

The target variable, namely the skill level, is a categorical feature. In the Malaysian context, skill level can be classified into skilled, semi-skilled and low-skilled categories (Department of Statistics Malaysia, 2025; Ministry of Human Resources, 2020). The nine major occupational groups in the Malaysia Standard Classification of Occupations (MASCO) are used to disaggregate labour into three skill level categories as depicted in Table 1. The definition of skill level is determined based on a single feature, namely, the occupational classification based on MASCO 2020 (Department of Statistics Malaysia, 2025). As a result, the current level of skill categorisation before model training is not influenced by other input predictors, including the income level. The income level is defined as an outcome measure that indicates the labour market returns to skill and is not included as an element in the classification of skill level categories (Zhao et al., 2025).

According to Figure 2, the semi-skilled category registered the highest cases at 74,717 with a percentage share of 62.00%. This was followed by the skilled category with 30,969 (25.70%) and the low-skilled category at 14,832 cases (12.31%). The skill level composition faced an imbalance in class structure, where classification models had a prediction preference for the majority category, causing lower prediction accuracy for the minority categories (Zhao & Li, 2025).

Data Preprocessing

A data preprocessing procedure was required, which included data cleaning, data transformation, data partitioning and handling imbalanced classes within the target variable. Data cleaning contained four procedures, namely eliminating duplicate cases, managing

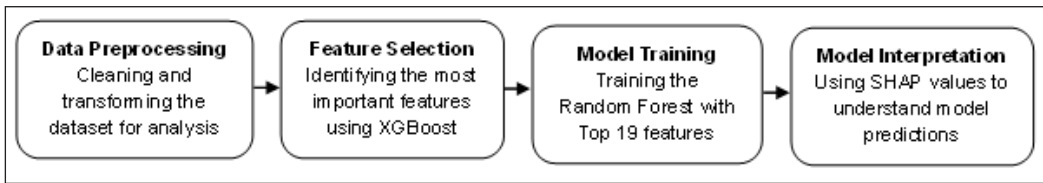


Figure 1. Research flow illustrating data preprocessing, feature selection, and model training and model interpretation

Table 1
Skill level categories based on major groups of MASCO 2020

Skill level categories	Major occupational group
Skilled	1: Managers
	2: Professionals
	3: Technicians and Associate Professionals
Semi-skilled	4: Clerical Support Workers
	5: Service and Sales Workers
	6: Skilled Agricultural, Forestry, Livestock and Fishery Workers
	7: Craft and Related Trades Workers
	8: Plant and Machine Operators and Assemblers
Low-skilled	9: Elementary Occupations

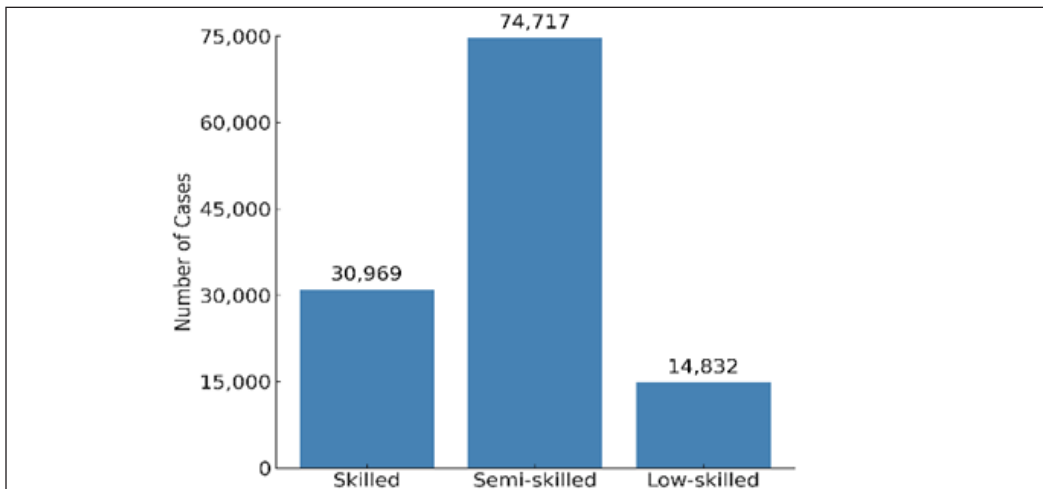


Figure 2. Distribution of skill level by skilled, semi-skilled and low-skilled categories

missing values, detecting outliers and eliminating irrelevant features. The dataset was randomly split into a ratio of 80:20, where 80 per cent of observations were assigned for training and 20 per cent for testing (Khalfallah et al., 2025; Sakyi-Yeboah et al., 2025) before the execution of the machine learning technique (Shaha et al., 2024).

An imbalanced dataset may limit decision-making validity, as it introduces bias in machine learning models towards majority categories while reducing the effectiveness of predictive performance for minority categories (Sheokand & Beniwal, 2025). Two of the most frequent practices used to handle imbalanced datasets are oversampling and undersampling. Oversampling generates synthetic data for minority categories while undersampling reduces the cases in majority categories, resulting in a similar distribution across all categories (Unlu & Subasi, 2025). In the meantime, undersampling may lead to the loss of information on the dominant category.

To address the multiclass imbalance classification problem within the target variable, the Synthetic Minority Oversampling Technique (SMOTE), one of the most prominent methods, was employed to ensure an equal distribution among the three skill level classes, namely skilled, semi-skilled and low-skilled, by generating synthetic cases for the minority classes (Unlu & Subasi, 2025). This finding is consistent with Tsai et al. (2024) who reported that employing SMOTE is effective in addressing multiclass imbalanced datasets.

In this study, the SMOTE was applied specifically to the training dataset to balance the samples (Wang et al., 2025). In labour market data, minority categories like low-skilled may represent a structurally different worker population instead of having a smaller number of samples due to undersampling. As a result, SMOTE was applied only to the training dataset to enhance classification performance. The testing dataset remained unchanged to maintain the actual labour market distribution.

Figure 3 illustrates the comparison of skill level distributions before and after implementing SMOTE. The original distribution indicated an imbalanced class with the semi-skilled component dominating the dataset at 74,717 cases, representing 62 per cent of the overall dataset. A dominant category of semi-skilled individuals resulted in high accuracy of prediction on the dominant class and low accuracy of prediction on the skilled and low-skilled individuals (Hasan et al., 2025). After the SMOTE technique implementation, these three categories obtained a balanced distribution with 58,099 cases for each category. This procedure was to ensure that the prediction model was not biased towards the majority class and to strengthen the model's effectiveness in correctly classifying minority skill level categories (Park et al., 2024). Further analysis of classification performance was conducted after addressing class imbalance using the SMOTE technique.

Subsequently, a prediction model was developed using Random Forest, as it was one of the most prominent and powerful ensemble machine learning algorithms. This bagging algorithm was integrated with SHAP to determine important features influencing skill level

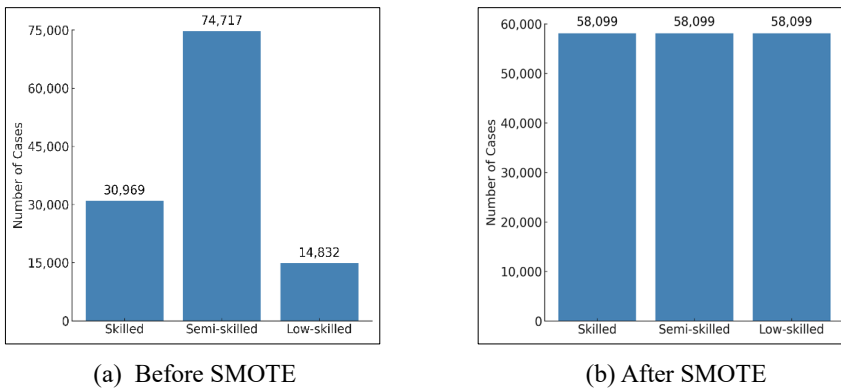


Figure 3. Distribution of skill level categories before and after applying SMOTE, demonstrating the imbalanced dataset into equal representation between skilled, semi-skilled and low-skilled workers

classification, thereby providing an explanation of the factors distinguishing the skilled, semi-skilled and low-skilled categories. A further explanation pertaining to the Random Forest algorithm and the SHAP interpretability was discussed in the subsequent segment of Model Training using Random Forest and Model Interpretation using Explainable SHAP analysis.

Feature Selection

Feature selection is an essential procedure for determining quality features in developing a reliable and strong skill level classification model (Hasan et al., 2025). These methods are categorised into filter, wrapper and embedded techniques. This process specified the relationship between input features and the target variables (Pedersen et al., 2025). It makes the selection of the most relevant input predictors that contribute to the skill level model prediction while removing the irrelevant features (Hasan et al., 2025).

Tree-based algorithms like Random Forest and XGBoost have built-in mechanisms for measuring the importance of features, which improve the efficiency in analysing relevant predictors in a classification model (Méndez-Astudillo, 2024). The main difference between the feature importance computation in Random Forest and XGBoost, as stated by Méndez-Astudillo (2024), is that Random Forest uses the total impurity reduction while XGBoost uses the average gain of the tree splits. However, XGBoost strength is when dealing with large volumes of data, various input predictors and determining the important features (Danquah et al., 2025; Méndez-Astudillo, 2024).

Based on the advantage of feature importance offered by XGBoost, a built-in feature selection method was employed to determine the most relevant input predictors to classify the skill level (Pedersen et al., 2025). XGBoost computes the feature importance score

using the gain method during its boosting process (Wang et al., 2024). According to Wang et al. (2024), XGBoost computes the feature importance score naturally in the training process and ranks nineteen features according to their contribution value. Empirical data presented by Wang et al. (2024) show that built-in feature importance is effective when dealing with large datasets, where there are multiple input predictors.

Accordingly, the XGBoost and Random Forest algorithms utilise different ensemble structures where boosting reduces bias and bagging reduces variance. However, both share similar inductive biases as they rely on decision tree-based learners. In line with the findings by Kurniadi and Saputri (2024), the combination of XGBoost and Random Forest leverages the high predictive accuracy and feature selection effectiveness of XGBoost, with the stability and robustness of Random Forest.

In accordance with the statement by Sariman et al. (2024), feature selection reduces the inclusion of irrelevant input features, which minimises the complexity of the model, but does not affect the skill level classification performance. As a consequence, the nineteen features ranked by built-in XGBoost were added to a Random Forest algorithm to improve the accuracy of skill level classification (Pourdaryaei et al., 2024). The deployment was carried out in nineteen iterations, starting with the highest importance score and proceeding to the least important score. These nineteen iterations were then evaluated through performance metrics to determine the proposed model. The following segment describes the Random Forest procedure and evaluation on these ranked input predictors.

Model Training using Random Forest

The skill level classification prediction used in this study applied the proposed Random Forest algorithm, which has been recognised due to its robust performance compared to other tree-based algorithms (Visibelli et al., 2025). This ensemble machine learning algorithm, based on the bagging technique, was introduced by Breiman (2001) and can be implemented to perform both classification and regression problems (Bodilsen et al., 2023; Liu & Wang, 2025). The algorithm constructs multiple decision trees separately using bootstrap samples of the original dataset. Those trees are trained on a random subset of input predictors at each split.

The randomness mechanism minimises the correlation between the single trees and increases the model generalisation, particularly when dealing with a large sample size, noisy data and high-dimensional datasets (Liu & Wang, 2025). The final Random Forest classification outcome is constructed by combining all the trees' predictions using a majority voting process, which helps to enhance accuracy and stability (Elshewey et al., 2025; Sakyi-Yeboah et al., 2025; Visibelli et al., 2025). This ensemble structure generates several decision trees with different split criteria, enabling stronger classification performance and removing irrelevant predictions (Salah-Ud-Din et al., 2024).

Random Forest is known for its effectiveness in handling complex datasets and robustness against overfitting (Visibelli et al., 2025). Additionally, this algorithm demonstrates strong generalisation, requires minimal data preprocessing procedures and achieves greater accuracy against the traditional algorithms even with smaller datasets (Zahra et al., 2025).

In this study, the Random Forest was implemented with the number of decision tree fixed at 100. The minimum sample size allowed to split an internal node was two, while the minimum number of samples at a leaf node was one. The Gini index was used as the splitting criterion to measure node impurity (Zeffora & Shobarani, 2022). This is due to its computational efficiency and is widely used in classification problems. There was no limit required on the maximum depth of the trees or even the number of leaf nodes, which allows the model to grow at full capacity and learn the complex decision boundaries within the skill level dataset. The random state was fixed at 42 to ensure reproducibility of the results (Hsieh, 2025). The set of parameters used was selected to balance the flexibility of the model and predictive power in the classification of skill level in the context of the Malaysian labour market.

Model Interpretation using Explainable SHAP Analysis

Interpretability techniques are an instrument to distinguish the underlying mechanisms of machine learning algorithms (Liu & Wang, 2025). It could be implemented before or during the model training (Neubauer et al., 2025). XAI approaches were developed to provide clarification into the operations of the models to facilitate end-user understanding (Neubauer et al., 2025). Those methods provided model transparency, strengthening trust in the predictive procedure and revealing the relationships between the target variable and the predictors (Liu & Wang, 2025). Leveraging XAI techniques was for the purpose of obtaining deeper explanations of the impacted features contributing to the performance of Random Forest algorithms in skill level classification. The most well-known model-agnostic approaches for interpreting the prediction mechanisms are SHAP, LIME, and Integrated Gradients (Joshi et al., 2025).

One of the prominent XAI techniques was SHAP analysis, introduced in 2017 by Lundberg and Lee (2017). The computation was an extension of the Shapley value measurement developed by Shapley (1953). The concept of SHAP was based on cooperative game theory, where each individual had a different contribution, and that allocation was distributed among the players following the proportion of their contributions (Liu & Wang, 2025). It relies on Shapley values to illustrate the feature's importance at both the global and local levels (Pedersen et al., 2025). SHAP supported interpretability at the global level by revealing a comprehensive understanding of the overall model behaviour, whereas the local level provided contribution scores of individual features (Dossche et al., 2025).

Considering these approaches, SHAP has been particularly prominent because of its transparency, reliability, and ability to explain model predictions according to individual characteristics using Shapley values (Adalarasu et al., 2025; Arravalli et al., 2025). Unlike traditional regression methods that provide only average effects through coefficients, SHAP illustrated how each feature contributes positively or negatively to individual predictions (Sani et al., 2025). The computation of SHAP values can be performed via interventional algorithms (Tree SHAP, Linear SHAP, Kernel SHAP) or conditional extensions of Kernel SHAP (Adalarasu et al., 2025). This framework supported both global explanations, summarising overall feature importance and local explanations, clarifying feature impacts for specific classes in multiclass classification tasks. Such interpretability is needed to comprehend the predictors of skill level classification as well as to facilitate evidence-based guidance on human capital initiatives.

Global SHAP provided the overall view of the skill level model predictions. The SHAP values described the importance of input features by providing a contribution score for individual features that reflected their impact on the predicted skill level categories (Ogungbire & Pulugurtha, 2025). The formula of SHAP was given in Equation 1 (Ogungbire & Pulugurtha, 2025).

$$\phi_i = \sum_{S \in F_{\{i\}}} \frac{|S|! (|F| - |S| - 1)!}{|F|!} [f_{S \cup \{i\}}(X_{S \cup \{i\}}) - f_S(X_S)] \quad [1]$$

where ϕ refers to the SHAP value of i^{th} feature, F refers to the total input features, S refers to the feature subsets excluding the i^{th} feature, $f_{S \cup \{i\}}$ and f_S refers to the model prediction without and with i^{th} feature, X_S refers to the input values of features for the subset S . In terms of multiclass classification, a positive number of SHAP values described the feature contribution to drive the classifier towards the exact category (Elshewey et al., 2025). On the contrary, a negative SHAP value showed the impact of a feature that influenced the prediction towards choosing a negative category (Elshewey et al., 2025). Additionally, a study by Sakyi-Yeboah et al. (2025) indicates that greater SHAP values reflected the level of importance of feature contribution towards the model classifier.

Local SHAP explained insights on the contribution of features to the specific skill level classes. The visualisation of SHAP values, highlighting the importance of features, was presented using dotted blue and red colours (Kaur et al., 2025). The effect of features could be analysed through red dots, which described a positive contribution, while blue dots indicated a negative contribution to the model prediction (Kaur et al., 2025; Ogungbire & Pulugurtha, 2025). Computing SHAP values underlined the most important features along with the impact and direction by offering an in-depth perspective on the mechanism of the model behaviour.

To overcome the problem of attribution inflation between correlated predictors such as education level and highest certificate, this study applies Tree SHAP by estimating conditional Shapley values based on the model outlined by Berge et al. (2024). This method explicitly considers inter-feature dependencies, whereby the observed values of the remaining input predictors are conditioned during the explanatory procedure. As a result, the model distributes contributions more fairly among related predictors and reduces the likelihood of inflating feature importance scores, a situation in which features are assumed to be independent.

Even though SHAP is computationally intensive, regular use by government bodies can be enabled through the implementation of less computationally intensive procedures, optimised algorithms and sampling methods (Mitchell et al., 2022). These strategies can reduce computational requirements and make SHAP a feasible solution for agencies with limited analytical resources. Furthermore, SHAP does not require extensive infrastructure, as this is not a real-time dataset, whereby the analysis is not needed frequently and does not require high-end infrastructure. The study offers some important insights, which not only enhance the accuracy of prediction but also offer a clear indication that could be used in human capital interventions and employer strategies.

Model Evaluation

The effectiveness of the machine learning algorithm was evaluated using classification metrics, including accuracy, sensitivity, specificity and F1-score (Bağcı & Kaya Soylu, 2025; Wang et al., 2024). In relation to multiclass classification, Bağcı and Kaya Soylu (2025) and Wang et al. (2024) indicates that a confusion matrix was implemented to determine true predictions involving true positives (TP) and true negatives (TN), whereas false predictions were defined as false positives (FP) and false negatives (FN).

The formula shown in Equation 2 specified the accuracy, quantified as the ratio of true predictions generated by the model to the total number of classification predictions (Rafrastara et al., 2025). In this context, accuracy reflected the ability of the Random Forest algorithm to effectively categorise workers into skilled, semi-skilled and low-skilled levels.

$$Accuracy = \frac{(TP1 + TP2 + TP3) + (TN1 + TN2 + TN3)}{(TP1 + TP2 + TP3) + (FP1 + FP2 + FP3) + (TN1 + TN2 + TN3) + (FN1 + FN2 + FN3)} \quad [2]$$

The F1-score was described as a combined metric between recall and precision, measured through its harmonic mean as computed in Equation 3 (Bağcı & Kaya Soylu, 2025). It described the model's capability to accurately determine the worker's skill level classification while minimising misclassification across the three skill level categories.

$$F1\ score = \frac{2 * Precision * Recall}{(Precision + Recall)} \quad [3]$$

Meanwhile, sensitivity measured the effectiveness of the classification model in capturing true positive cases for each skill level category, as given in Equation 4 (Ul Haq et al., 2024). It examined how effectively the model recognised the worker's actual skill level category.

$$Sensitivity = \frac{(TP1 + TP2 + TP3)}{(TP1 + TP2 + TP3) + (FN1 + FN2 + FN3)} \quad [4]$$

In this study, specificity was examined by the ability of the classification model to distinguish true negative cases among the three skill level categories, which were defined as Equation 5 (Ul Haq et al., 2024). It prevented workers from being wrongly classified into other skill level categories beyond their actual skill category.

$$Specificity = \frac{(TN1 + TN2 + TN3)}{(TN1 + TN2 + TN3) + (FP1 + FP2 + FP3)} \quad [5]$$

RESULTS AND DISCUSSION

This section highlights three primary outcomes. It begins with the descriptive analysis of the 2023 Salaries and Wages dataset, followed by the feature selection outcomes based on XGBoost feature importance, Random Forest classification performance and interpretability of XAI using the SHAP technique.

Descriptive Analysis

The Salaries and Wages Survey for the reference year 2023 comprised 120,518 workers. The illustration, as shown in Figure 4, emphasises the composition of educational attainment, with the majority having completed secondary education at 54.7 per cent, followed by tertiary education (33.9 per cent). Only a small segment had completed primary education at 9.3 per cent, whereas a relatively minimal share of 2.2 per cent reported having no formal education. This finding describes an educated workforce recommending potential and readiness for greater skill enhancement initiatives aimed at increasing skilled labour in the country.

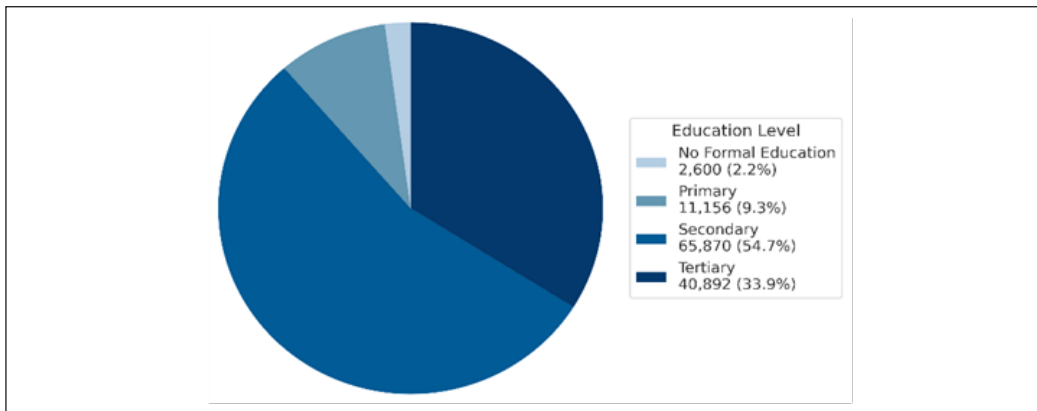


Figure 4. Composition of educational attainment

XGBoost Feature Importance

The XGBoost feature importance analysis indicated the importance scores for nineteen features in the dataset, as depicted in Figure 5. It demonstrated that education level, with a score of 0.213, was the most relevant predictor to determine skill level categories. Salaries and wages ranked second (0.146), followed by the highest certificate (0.118). Concurrently, employment status (0.095) and activity (0.067) also played important roles during the model training. The demographic and labour market characteristics, including strata, gender, hours worked and marital status, had a moderate influence on the skill level categories. Features such as district and disability status had minimal effects, whereas citizenship did not contribute to the skill level prediction. These findings confirm that labour market indicators, particularly education, salaries and wages, and activity are the primary drivers in the model for further SHAP interpretability analysis.

Selected feature importance rankings have been utilised in other labour market studies. Jamaludin et al. (2022) highlighted that salaries and wages were an important indicator as they influenced labour mismatch in the labour market scenario. According to Mansor and Othman (2024), education was critical as the higher level of education influenced the employability in the labour market demand. From the unemployment perspective, Celbiş (2023) indicated that age and education influenced the opportunities for an individual to participate in the labour market.

Despite the dataset being from the 2023 reference year, it shows stability because the statistics reflect the underlying structural dynamics of the Malaysian labour market. Macroeconomic shocks can affect the volume of employment (Department of Statistics Malaysia, 2025), but they do not cause immediate changes in entry-level, including education level and the highest certificate, which are prerequisites for a specific skill level (Ministry of Human Resources, 2020). Additionally, the Thirteenth Malaysia Plan

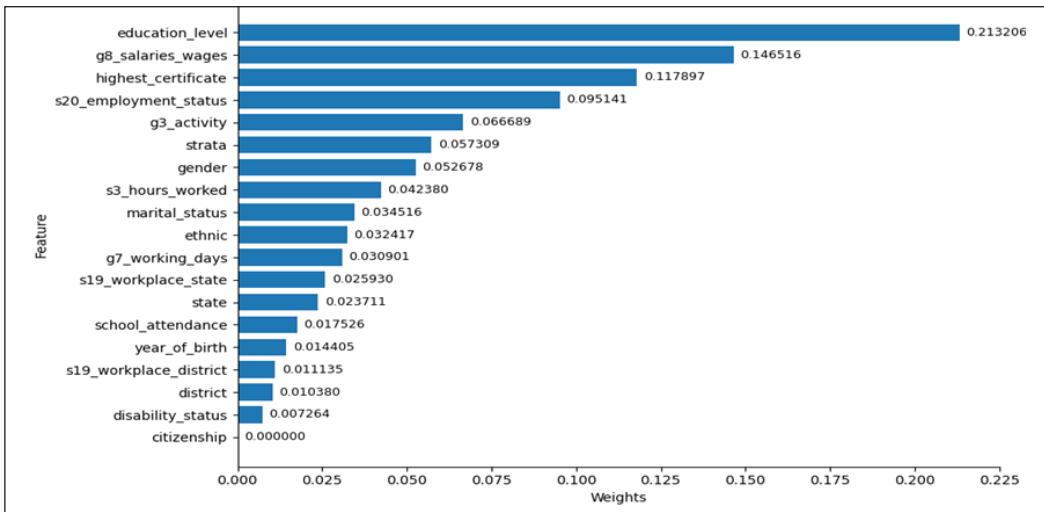


Figure 5. List of XGBoost Feature Importance based on nineteen features

2026-2030 indicates that challenges such as skill mismatch and concentration in low-skill occupations have been persistent issues over a long period, thus confirming the stability of these trends (Ministry of Economy, 2025). Consequently, the findings reflect structural association patterns at one reference period and should not be interpreted as causal effects between skill level categories.

This study employed XGBoost feature importance ranking due to its computational efficiency and suitability for large datasets (Danquah et al., 2025). Different feature selection techniques may produce different feature subsets and interpretation outcomes (Wang et al., 2024). This study considers skill level as a fixed classification based on cross-sectional data. It implicitly assumes limited short-term mobility across skill level categories. The use of longitudinal data for further studies would allow skill development analysis as a dynamic trajectory.

Random Forest Performance

Despite the skill level being determined by MASCO occupational guidelines, this study does not attempt to reproduce these guidelines. Instead, it aims to estimate the association between socioeconomic characteristics and skill level categories. The predictive model identifies the most important determinants of skill level status and assigns weights to their relative importance. It also enables estimation when occupational information is incomplete. For instance, the model can predict skill level when occupational information is not provided in administrative datasets. As a result, the model provides explanatory and policy-relevant insights that extend beyond the fixed MASCO classification.

This study employed a powerful machine learning technique to operationalise this predictive framework and capture the non-linear relationships among the input predictors.

To verify the effectiveness of the Random Forest classification model, the performance was examined in previous studies. In the prior literature, a comparable analysis of five tree-based algorithms consisting of Decision Tree, Random Forest, Gradient Boosting Machines, Adaptive Boosting and Extreme Gradient Boosting was examined for multiclass skill level prediction (Rabi'atul'adawiah Shablia et al., 2025). The results demonstrated that Random Forest recorded an outstanding predictive performance across accuracy (0.8645), specificity (0.9089), sensitivity (0.8645) and F1-score (0.8636) when dealing with imbalanced skill level categories. Thus, the combination of Random Forest with SHAP explanations was further investigated to enhance transparency and interpretability.

This study determined Random Forest as the most effective classification model for classifying skill level categories in the Malaysian labour market. The findings were consistent with prior classification studies across several domains, including medical and other labour market studies, where Random Forest consistently outperformed other tree-based algorithms (Arravalli et al., 2025; Bağcı & Kaya Soylu, 2025; Khalfallah et al., 2025; Rafrastara et al., 2025; Visibelli et al., 2025). Among others were predictions of therapy dropout among chronic pain patients (Visibelli et al., 2025), detection of breast cancer symptoms (Arravalli et al., 2025), fertility preferences among Somalian women (Sani et al., 2025) and employee attrition (Manafi Varkiani et al., 2025). Taken together, these studies suggest that the robust performance of Random Forest was mainly attributed to its ensemble structure (Khalfallah et al., 2025), efficiency in handling feature redundancy, and the ability to recognise complex classification patterns in the dataset (Lalithadevi & Krishnaveni, 2025) and the ability to mitigate overfitting (Visibelli et al., 2025).

Thus, in this study, the Random Forest classifier was constructed using the top nineteen features obtained from XGBoost feature importance. The selection of sixteen features was made when the predictive accuracy had reached its stability, as illustrated in Figure 6. To provide a statistical rationale for the selected cut-off, rather than relying on visual inspection, a formal stopping criterion was employed, since adding more features to the model produced only minimal increases in accuracy (Syeda & Asghar, 2023). As highlighted by Kempny et al. (2025), there is no common cut-off threshold that is reliable across datasets for defining the stopping criterion. This demonstrates that incorporating additional features only provides a marginal performance improvement while increasing model complexity. This exhibits a balance between model parsimony and sufficient explanatory information.

The feature subset comprised socioeconomic, demographic and employment-related characteristics. These include education attainment, salaries and wages, highest certificate, employment status, activity, strata, gender, hours worked, marital status, ethnic, working days, workplace state, state, school attendance, year of birth and workplace district.

Utilising multidimensional predictors assisted the algorithm's ability to recognise the patterns and relationships affecting skill level categorisation. The effectiveness of this algorithm was further assessed through model evaluation.

To validate that SMOTE did not distort the underlying structure of the training dataset, the distributions of the primary input predictors were compared between the original training dataset and the SMOTE dataset (Mujahid et al., 2024). Continuous features, including salaries and wages hours worked, and year of birth, showed similar patterns in the kernel density plots. For the categorical features, including education level and activity, only slight differences were observed. No artificial clusters or unrealistic values were found. These findings suggest that SMOTE enhanced class balance without changing the heterogeneity of the original data (Mujahid et al., 2024).

To ensure that the predictive model is not driven primarily by salaries and wages, an additional analysis was conducted by removing the salaries and wages feature from the model. The classification accuracy decreased slightly from 0.8645 to 0.8334 (Figure 6). This emphasises that although salary and wages are informative, the overall model performance does not change substantially when the feature is excluded. Accordingly, the predictive outcome is not driven solely by salaries and wages as other socioeconomic and employment characteristics also play an important role in defining the skill level categories. These findings are consistent with evidence from South Korean where skill level does not always improve employability or salaries, indicating that labour market performance, such as salaries, is not necessarily determined solely by individual skills (Ban & Choi, 2025).

After the robustness checks, the predictive performance is summarised through a confusion matrix as exhibited in Figure 7(a). This multiclass classification was further examined by using four evaluation metrics of accuracy, sensitivity, specificity and F1-score, as illustrated in Figure 7(b). The accuracy score of the proposed algorithm was 0.8754, which indicates that most predictions were in concordance with the appropriate skill level categories. As specified by Halias et al. (2023), an accuracy of more than 70 per cent is typically considered to indicate high classification performance.

The F1-score of 0.8378 reflects a balance between recall and precision, which demonstrates Random Forest's strong performance in all categories. The high sensitivity score at 0.8485 illustrates the ability of the Random Forest algorithm to classify each category correctly. A scoring of 0.9384 for specificity reveals the effectiveness of the algorithm in reducing the misclassification of skill level categories. Collectively, these findings suggest that Random Forest is efficient at classifying skill level into the three categories and is suitable for multiclass classification prediction tasks involving skill level. Based on the confusion matrix, most misclassifications occur at the boundaries between categories, particularly between the semi-skilled category and its neighbouring categories.

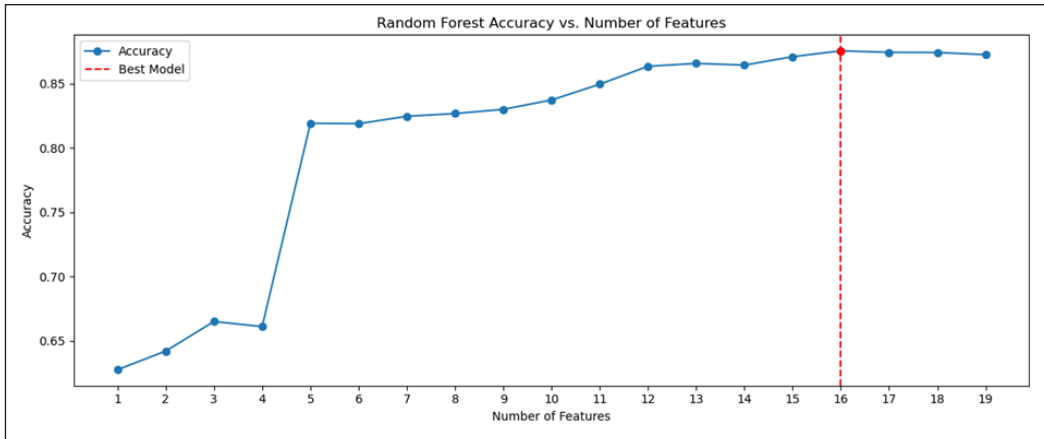


Figure 6. The prediction accuracy of the Random Forest algorithm across various feature subsets in the testing dataset. The red dotted line portrayed the optimum Random Forest algorithm with sixteen features, which had the highest accuracy score

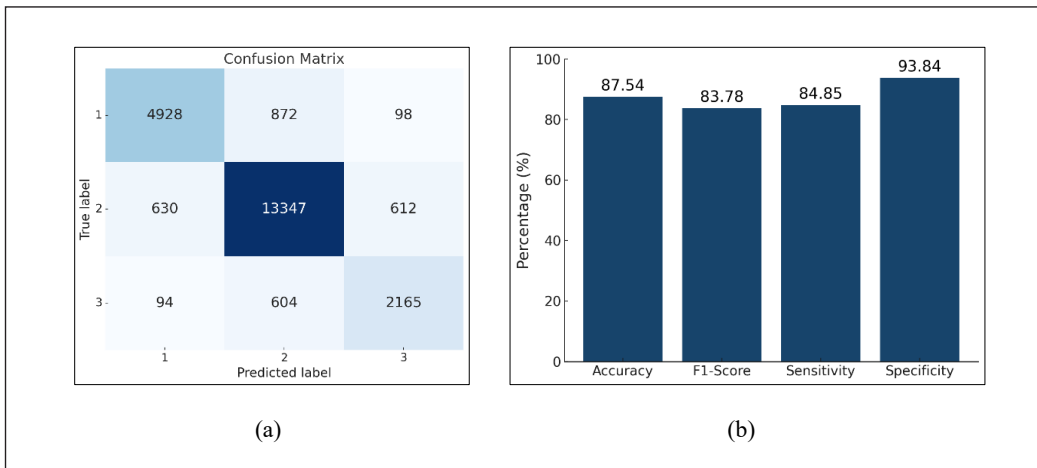


Figure 7. Random Forest classification results (a) Top 16 features confusion matrix and (b) Model evaluation of the proposed algorithm.

This reflects the nature of the MASCO occupational taxonomy, where clearly defined categories usually share similar skill profiles, rather than a weakness of the predictive model (Ministry of Human Resources, 2020).

The predictive performance for the three skill level categories is presented in Table 2. For the skilled category, the model demonstrates strong classification capabilities with an accuracy of 0.9275. The high sensitivity of 0.8719 and specificity of 0.9452 resulted in an F1-score of 0.8533.

Table 2
Predictive performance by skill level categories

Category	Accuracy	Sensitivity	Specificity	F1-score
Skilled	0.9275	0.8719	0.9452	0.8533
Semi-skilled	0.8836	0.9004	0.8543	0.9076
Low-skilled	0.9397	0.7530	0.9659	0.7546
Skill Level	0.8754	0.8485	0.9384	0.8378

This reflects that the model can distinguish skilled workers from other categories with minimal false positives. In the case of the semi-skilled category, the model registered an accuracy of 0.8836. The higher sensitivity of 0.9004 shows strong predictive ability for this category. However, the lower specificity of 0.8543 indicates a moderate false positive rate, where workers from other categories are classified as semi-skilled. For the minority category, the low-skilled category, the model achieved the highest accuracy of 0.9397 and specificity of 0.9659. This is largely due to the correct rejection of skilled and semi-skilled categories, the true negative cases. On the contrary, sensitivity decreased to 0.7530, resulting in a lower F1-score of 0.7546. This finding shows that the model can prevent false classification of other categories as low-skilled, but it is prone to missing actual low-skilled workers.

To evaluate sensitivity to hyperparameter choices, the Random Forest algorithm was estimated using alternative tree structures. These included shallow trees, which often provide good generalisation with a maximum depth of 10 (Zhou & Mentch, 2023) and regularised trees that prevent overfitting with a maximum depth of 20 and a minimum leaf size of five (Hancock & Khoshgoftaar, 2022). The same SMOTE dataset and validation dataset were utilised for all configurations. The baseline setting achieved the best performance with an accuracy of 0.8754 and an F1-score of 0.8378. Similar results were observed with moderate regularisation, where the accuracy was 0.8391. However, when using heavily constrained shallow trees, the accuracy decreased to 0.7774. The feature importance rankings remained stable across configurations with a correlation value of 0.9828, indicating consistent interpretation. These findings demonstrate that the predictive performance and SHAP explanations remain strong when the hyperparameter settings are slightly changed and do not require extensive tuning. The following segment explains how SHAP was applied to interpret the Random Forest classification model.

The Interpretability of SHAP Analysis

Although the goal is to obtain high accuracy, interpretability is essential in making policy decisions. The XAI technique, particularly SHAP analysis, was employed to interpret the impact of features on the algorithm while enhancing the model transparency (Arravalli et

al., 2025; Visibelli et al., 2025). Despite offering valuable information, the implementation of SHAP analysis requires high computational resources (Joshi et al., 2025). Integration of SHAP analysis with the proposed Random Forest algorithm offered valuable information on the most important factors impacting skill level within the Malaysian labour market. Further analysis on the illustration of skill level can be described through red dots (positive outcome) and blue dots (negative outcome) as portrayed in Figure 8. The plot of SHAP values clearly portrayed the importance of the features for predicting the skill level classification. Greater SHAP values describe higher contributions of the specific features in influencing the classification model outcomes into a particular skill level category.

SHAP analysis determined the most influential factors for predicting skill level. The summary from the SHAP global perspective presented in Figure 8(a) identified nine input predictors as the most impactful features in skill level classification. The subsets were salaries and wages activity, highest certificate, education level, working days, gender, employment status, strata and year of birth. Educational attainment represents the highest education level completed by an individual, comprising no formal education, primary, secondary and tertiary level. Meanwhile, the highest certificate is closely related but refers to specific formal qualifications, which consist of no certificate/ not applicable, SPM and below, STPM/ certificate, diploma and degree (Department of Statistics Malaysia, 2025).

Salaries and wages were demonstrated as the most dominant predictor that was strongly associated with higher skill level prediction. This pattern is consistent with the labour market scenario where skilled occupations are often associated with higher remuneration (Zhao et al., 2025). The dominance of salaries and wages in SHAP interpretability may reflect institutional and job allocation mechanisms in the labour market, rather than individual skill differences alone (Zhao et al., 2025). As a result, salaries and wages are interpreted as indicators of labour market performance instead of as direct measures of individual skill, consistent with the objective of this study.

Likewise, the type of activity performed by workers also impacted the prediction of skill level, as only certain sectors are demanding skilled workers. In the meantime, the highest certificate and education level were associated with discovering the type of skill level. It was due to skilled workers who had undergone formal education and academic criteria, which is important for career development. The findings of this study are consistent with Koltai et al. (2025), who highlighted through principal component analysis that the education level, income and occupational prestige are essential indicators in socio-economic studies, regardless of gender and age. Another five influencing predictors included working days, gender, employment status, strata and the year of birth.

Despite education level, highest certificate and salaries and wages being conceptually related, tree-based algorithms such as Random Forest are less sensitive to multicollinearity, as Random Forest can tolerate correlated input predictors and maintain its stable

classification performance (Soni et al., 2025). The SHAP values represent the conditional impact of these features on the model and do not indicate independent causal relationships. The interpretation focuses on their overall explanatory role in classifying the skill level and not on their individual effects.

Although SHAP identifies several demographic features like gender and year of birth as statistically related to the model predictions, they do not directly cause skill level classification. However, these input predictors reflect broader labour market conditions, including differences across the age groups and types of jobs. To reduce the risk of misuse in sensitive policy areas, the findings presented in this study are strictly limited to macro-level information intended for strategic human capital planning. According to the Thirteenth Malaysia Plan 2026-2030 document, human capital development is designed to support the national agenda rather than individual-level decisions (Ministry of Economy, 2025). Therefore, the findings should not be applied to individual evaluation or unfair decision processes. This is consistent with the nature of SHAP explanations in emphasising patterns in model behaviour rather than causal relationships (Joshi et al., 2025). SHAP is employed in this study as a tool to assist in explaining the model behaviour, particularly on the contribution of individual input predictors to the skill level classification (Joshi et al., 2025; Neubauer et al., 2025). Meanwhile, the economic interpretation should not rely on feature importance alone. The SHAP explanations analysed in this study are intended to support policy discussion by emphasising observed patterns in skill level model behaviour, but not to imply causal relationships for policy initiatives (Neubauer et al., 2025).

Exploring the local SHAP analysis revealed the variations across those three skill level categories. The SHAP summary showed that the most relevant predictors influencing skilled and semi-skilled categories were largely the same features. Nevertheless, the importance score varies depending on the SHAP values. As presented in Figure 8(b), the top three highest SHAP values indicated skilled category predictors were salaries and wages highest certificate and activity, whereas the semi-skilled in Figure 8(c) demonstrated highest certificate, salaries and wages and activity in determining the worker's skill. The outcomes were consistent with Jamaludin et al. (2022), which indicated salaries and wages as the primary indicators of labour market mismatch. In addition to income factors, educational attainment also played a vital role as it impacted graduate employability performance and the labour market requirements (Mansor & Othman, 2024). The activity performed by the worker was identified as one of the relevant features in occupational injury studies (Zermane et al., 2023).

Unlike the skilled and semi-skilled categories, Figure 8(d) shows that the classification of low-skilled workers did not rely on working days, year of birth and strata but rather on marital status, state and workplace. A study carried out by Kaboth et al. (2023) showed that low-skilled occupations typically did not require any vocational training, and their

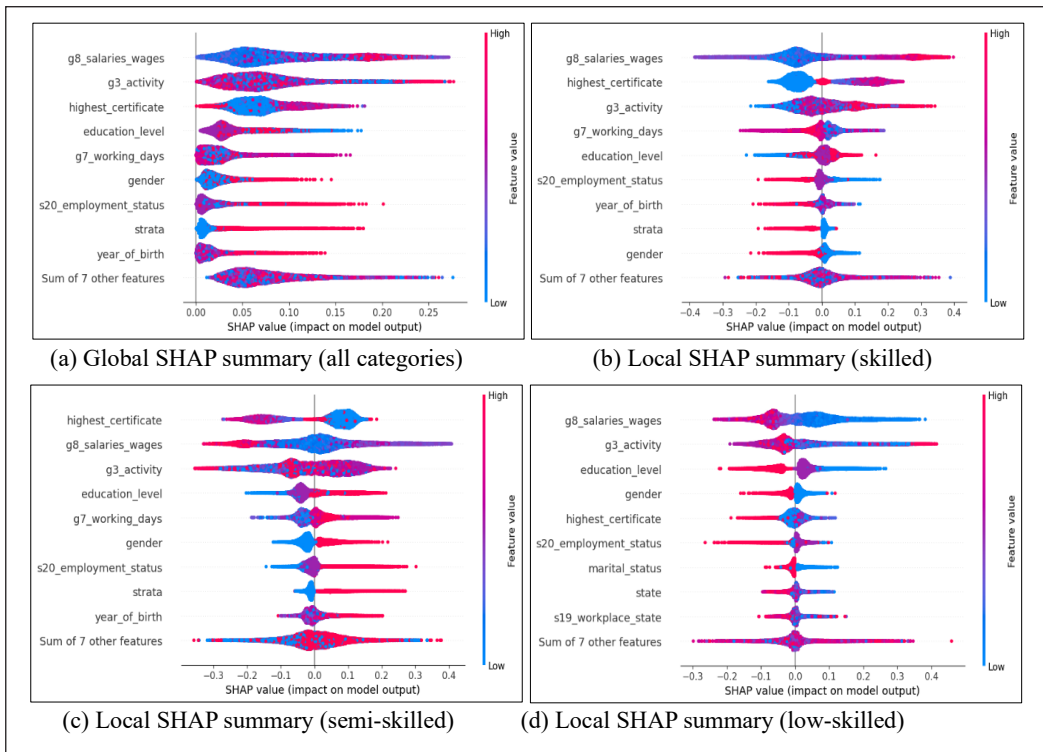


Figure 8. Summary of Global and Local SHAP (a) Global SHAP summary (all categories); (b) Local SHAP summary (skilled); (c) Local SHAP summary (semi-skilled); and (d) Local SHAP summary (low-skilled)

positions were frequently replaceable. This situation was attributed to the nature of low-skilled occupations, which typically require minimal training and experience to perform the required tasks (Kaboth et al., 2023).

The local SHAP analysis discloses how similar input predictors operate across the three skill level categories. The visualisation shows that salaries and wages received, and the highest certificate obtained are the most prominent predictors of the skilled category. Higher salaries and wages and higher certificates show a stronger positive impact on the model, demonstrating that they push the classification model towards categorising workers as skilled. In the semi-skilled category, the highest certificate and activity are the key drivers of classification. The SHAP pattern assumes that semi-skilled workers are identified by specific job activities, together with an average level of academic qualifications. For the low-skilled category, the model largely relies on the salaries and wages and education level to make predictions. The findings show that earning lower salaries and wages and a lower educational level are the main characteristics and are associated with a higher likelihood of being classified as low-skilled.

Overall, these results indicate that SHAP analysis determined salaries and wages activity and highest certificate as the main predictors in predicting skill level categories, despite their importance scores varying across the categories. Taken together, the interpretation technique from two perspectives, including global and local levels, enhances model understanding and provides beneficial insights that can be applied to labour market policy.

CONCLUSION

This study develops an interpretable skill level classification model for the Malaysian labour market by integrating the Random Forest algorithm with XAI techniques. The results of this study indicate that with the use of the top sixteen important features, the Random Forest algorithm achieves a high predictive accuracy of 0.8754. It shows a strong performance in all the evaluation metrics and thus outperformed other tree-based algorithms. Predictive accuracy combined with the incorporation of SHAP analysis also increases the trustworthiness, transparency and interpretability of the proposed model. The SHAP values determine the most affecting features by measuring both positive and negative impacts, with salaries and wages being the most influential global predictors, followed by activity and highest certificate. SHAP explanations at the local level reveal the feature contributions for each category of skilled, semi-skilled and low-skilled, offering in-depth insights into the underlying causes of classification differences. As this study uses a single cross-sectional reference year dataset, the findings are reliable and applicable to situations where the economic structure and labour market characteristics are like those in 2023.

Based on these findings, the interpretable framework supports policymakers in comprehending the rationale and justification of the predicted outcomes and enables the identification of socioeconomic, demographic, and employment characteristics that shape the classification of skill level. This information promotes evidence-based justification of human capital development and is consistent with current labour market practices. SHAP visualisations should be used as a tool to support model interpretability and interpreted together with the labour market context. It is not appropriate to be analysed alone for decision-making purposes.

However, the framework still relies on MASCO occupational groupings. The list of occupations only appears in the Labour Force/ Salaries & Wages/ Migration Survey conducted by the Department of Statistics Malaysia. Other labour market information does not contain occupational data, and the skill level cannot be measured directly. This study, therefore, focuses on identifying relevant input predictors that are strongly associated with skill level. These predictors can be applied in future studies to predict skill level using secondary data that do not capture occupational information. As skill level in this study is defined using the MASCO occupational classification, the findings should be interpreted within this definition.

To enhance the validity of this study, future studies should adopt other feature selection methods, other advanced interpretability approaches, and time-series analyses to strengthen generalisability and reflect the transitions in the Malaysian labour market. To measure the utility of policy interventions, an extended study should use a longitudinal comparison framework to evaluate whether SHAP-based interventions yield better human capital outcomes and reduce skill mismatches compared to traditional approaches. This study proposes a new framework to identify important features for skill level classification in Malaysia, addressing current approaches that rely on a single feature. Taken together, this result suggests that the combination of the Random Forest algorithm and SHAP explanations enables an accurate and interpretable prediction of skill level classification by addressing the gap between data-driven insights and the practical application in the Malaysian labour market.

ACKNOWLEDGEMENT

The authors are very grateful to the Department of Statistics Malaysia for providing the annual labour market statistics. Heartfelt appreciation is also extended to Universiti Teknologi MARA (UiTM) and the University of Leeds for their support, recommendations and guidance in the research undertaking.

REFERENCES

- Adalarasu, K., Chetty, R. M. K., Begum, K. G., Harini, S., & Janardhanan, M. (2025). An explainable machine learning (XAI) framework for classification of intricate dancing posture among Indian Bharatanatyam dancers. *Applied Soft Computing*, 171, Article 112817. <https://doi.org/10.1016/j.asoc.2025.112817>
- Arravalli, T., Chadaga, K., Muralikrishna, H., Sampathila, N., Cenitta, D., Chadaga, R., & Swathi, K. S. (2025). Detection of breast cancer using machine learning and explainable artificial intelligence. *Scientific Reports*, 15(1), Article 26931. <https://doi.org/10.1038/s41598-025-12644-w>
- Bağcı, M., & Kaya Soylu, P. (2025). The optimal threshold selection for high-frequency pairs trading via supervised machine learning algorithms. *Computational Economics*. Advance online publication. <https://doi.org/10.1007/s10614-025-10958-5>
- Ban, G. W., & Choi, H. (2025). Skills and labour market performance in Korea: An international comparison perspective. *International Economic Journal*, 39(2), 357-382. <https://doi.org/10.1080/10168737.2025.2463372>
- Berge, H., Glad, I. K., Jullum, M., & Aas, K. (2024). A comparative study of methods for estimating model-agnostic Shapley value explanations. *Data Mining and Knowledge Discovery*. Advance online publication. <https://doi.org/10.1007/s10618-024-01016-z>
- Bodilsen, S., Albeck Nielsen, S., & Rosholm, M. (2023). Measuring employment readiness for hard-to-place individuals. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4646348>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>

- Celbiş, M. G. (2023). Unemployment in rural Europe: A machine learning perspective. *Applied Spatial Analysis and Policy*, 16(3), 1071-1095. <https://doi.org/10.1007/s12061-022-09464-0>
- Celbiş, M. G., Wong, P., Kourtit, K., & Nijkamp, P. (2023). Impacts of the COVID-19 outbreak on older-age cohorts in European labour markets: A machine learning exploration of vulnerable groups. *Regional Science Policy & Practice*, 15(3), 559-585. <https://doi.org/10.1111/rsp3.12520>
- Danquah, L. K. G., Appiah, S. Y., Mantey, V. A., Danlard, I., & Akowuah, E. K. (2025). Computationally efficient deep federated learning with optimised feature selection for IoT botnet attack detection. *Intelligent Systems with Applications*, 25, Article 200462. <https://doi.org/10.1016/j.iswa.2024.200462>
- Department of Statistics Malaysia. (2024a). *Labour force/salaries & wages/migration survey 2024*. <https://www.dosm.gov.my/portal-main/census-surveys-list/labour-force-social-statistics>
- Department of Statistics Malaysia. (2024b). *Salaries & wages survey 2023*. <https://www.dosm.gov.my/portal-main/release-content/salaries-and-wages-survey-report-2023>
- Department of Statistics Malaysia. (2025). *Salaries & wages survey report 2024*. <https://www.dosm.gov.my/portal-main/release-content/salaries-and-wages-survey-report-2024>
- Dhungana, H. (2025). A machine learning approach for wind turbine power forecasting for maintenance planning. *Energy Informatics*, 8(1), Article 2. <https://doi.org/10.1186/s42162-024-00459-4>
- Dossche, W., Vansteenkiste, S., Baesens, B., & Lemahieu, W. (2025). Anticipating delays in recruitment: Explainable machine learning for the prediction of hard-to-fill online job vacancies. *European Journal of Operational Research*. Advance online publication. <https://doi.org/10.1016/j.ejor.2025.06.027>
- Elshewey, A. M., Selem, E., & Abed, A. H. (2025). Improved CKD classification based on explainable artificial intelligence with extra trees and BBFS. *Scientific Reports*, 15(1), Article 17861. <https://doi.org/10.1038/s41598-025-02355-7>
- Halias, A. F., Saiful, N. H., Ibrahim, N., Muhamad Jamil, S. A., Mansor, M. M., Ul-Saufie, A. Z., & Md Ghani, N. A. (2023). Type 2 diabetes mellitus prediction using data mining approach. In *2023 IEEE International Conference on Computing (ICOCO)* (pp. 29-34). <https://doi.org/10.1109/ICOCO59262.2023.10398078>
- Hancock, J., & Khoshgoftaar, T. M. (2022). Optimising ensemble trees for big data healthcare fraud detection. In *Proceedings of the International Conference on Information Reuse and Integration for Data Science (IRI)* (pp. 65-72). <https://doi.org/10.1109/IRI54793.2022.00061>
- Hasan, Y., de Lima, A., Namjoo, E., de Bulnes, D. F., Albarracín, J. F. H., & Ryan, C. (2025). Improving breast cancer diagnosis using grammatical evolution-based feature selection. *SN Computer Science*, 6(4), Article 306. <https://doi.org/10.1007/s42979-025-03840-6>
- Hsieh, P.-H. (2025). ML-based validation of experimental randomisation in learning games. *Frontiers in Artificial Intelligence*, 8, Article 1541087. <https://doi.org/10.3389/frai.2025.1541087>
- Jamaludin, S., Safian, S. S., Said, R., & Muhamed, A. A. (2022). The incidence of overeducated and undereducated worker for manufacturing sector in Malaysia. *International Journal of Academic Research in Business and Social Sciences*, 12(5), 810-823. <https://doi.org/10.6007/IJARBSS/v12-i5/12942>
- Joshi, A., Pradhan, B., Chakraborty, S., Varatharajoo, R., Alamri, A., Gite, S., & Lee, C.-W. (2025). An explainable Bi-LSTM model for winter wheat yield prediction. *Frontiers in Plant Science*, 15, Article 1491493. <https://doi.org/10.3389/fpls.2024.1491493>

- Kaboth, A., Hünefeld, L., & Himmelreicher, R. (2023). Employment trajectories of workers in low-skilled jobs in Western Germany. *Journal for Labour Market Research*, 57(1), Article 21. <https://doi.org/10.1186/s12651-023-00351-w>
- Kaur, M., Valderrama, C. E., & Liu, Q. (2025). Hybrid vision transformer framework for efficient and explainable SEM image-based nanomaterial classification. *Machine Learning: Science and Technology*, 6(1), Article 015066. <https://doi.org/10.1088/2632-2153/adc072>
- Kempny, C., Annac, K., Wahidie, D., Yilmaz-Aslan, Y., & Brzoska, P. (2025). Validation of stop criteria in ASReview for optimising title and abstract screening. *European Journal of Public Health*, 35(Supplement_4). <https://doi.org/10.1093/eurpub/ckaf161.1276>
- Khalfallah, S., Puech, W., Tlija, M., & Bouallegue, K. (2025). Exploring the effectiveness of machine learning and deep learning techniques for EEG signal classification in neurological disorders. *IEEE Access*, 13, 17002-17015. <https://doi.org/10.1109/ACCESS.2025.3532515>
- Koltai, J., Rakovics, Z., Kmetty, Z., Számel, K., Ungvári, B., Váradi, B., & Huszár, Á. (2025). Classifying social position with social media behavioural data. *EPJ Data Science*, 14(1), Article 60. <https://doi.org/10.1140/epjds/s13688-025-00578-2>
- Kurniadi, F. I., & Saputri, H. A. (2024). Feature selection using XGBoost on METABRIC dataset for survivability breast cancer detection. In *2024 International Conference on Information Management and Technology (ICIMTech)* (pp. 259-262). <https://doi.org/10.1109/ICIMTech63123.2024.10780791>
- Lalithadevi, B., & Krishnaveni, S. (2025). ExAIRFC-GSDC: An advanced machine learning-based interpretable framework for accurate gas leakage detection and classification. *International Journal of Computational Intelligence Systems*, 18(1), Article 16. <https://doi.org/10.1007/s44196-025-00742-6>
- Lee, J.-Y., Lee, W., & Cho, S. (2023). Characteristics of fatal occupational injuries in migrant workers in South Korea: A machine learning study. *Heliyon*, 9(9), Article e20138. <https://doi.org/10.1016/j.heliyon.2023.e20138>
- Liu, J., & Wang, H. (2025). Impact of personal information and reputation system on gig workers' employment status: An interpretable machine learning-based approach. *Journal of Computational Social Science*, 8(3), Article 59. <https://doi.org/10.1007/s42001-025-00371-1>
- Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions (arXiv:1705.07874). *arXiv*. <https://doi.org/10.48550/arXiv.1705.07874>
- Manafi Varkiani, S., Pattarin, F., Fabbri, T., & Fantoni, G. (2025). Predicting employee attrition and explaining its determinants. *Expert Systems with Applications*, 272, Article 126575. <https://doi.org/10.1016/j.eswa.2025.126575>
- Mansor, A., & Othman, Z. (2024). Malaysian community college graduates' employability prediction model using a machine learning approach. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 16(3), 1-7. <https://doi.org/10.54554/jtec.2024.16.03.001>
- Méndez-Astudillo, J. (2024). The impact of comorbidities and economic inequality on COVID-19 mortality in Mexico: A machine learning approach. *Frontiers in Big Data*, 7, Article 1298029. <https://doi.org/10.3389/fdata.2024.1298029>

- Ministry of Economy. (2025). *Thirteenth Malaysia plan 2026-2030*
- Ministry of Human Resources. (2020). *Malaysia standard classification of occupations 2020*.
- Mitchell, R., Frank, E., & Holmes, G. (2022). GPUTreeShap: Massively parallel exact calculation of SHAP scores for tree ensembles. *PeerJ Computer Science*, 8, Article e880. <https://doi.org/10.7717/peerj-cs.880>
- Mujahid, M., Kina, E., Rustam, F., Villar, M. G., Alvarado, E. S., Diez, I. D. L. T., & Ashraf, I. (2024). Data oversampling and imbalanced datasets: An investigation of performance for machine learning and feature engineering. *Journal of Big Data*, 11(1), 1-32. <https://doi.org/10.1186/s40537-024-00943-4>
- Neubauer, A., Brandt, S., & Kriegel, M. (2025). Explainable multi-step heating load forecasting: Using SHAP values and temporal attention mechanisms for enhanced interpretability. *Energy and AI*, 20, Article 100480. <https://doi.org/10.1016/j.egyai.2025.100480>
- Ogungbire, A., & Pulugurtha, S. S. (2025). Spatiotemporal risk mapping of statewide weather-related traffic crashes: A machine learning approach. *Machine Learning with Applications*, 20, Article 100642. <https://doi.org/10.1016/j.mlwa.2025.100642>
- Park, J., Choi, S., & Feng, Y. (2024). Predicting startup success using two bias-free machine learning: Resolving data imbalance using generative adversarial networks. *Journal of Big Data*, 11(1), Article 122. <https://doi.org/10.1186/s40537-024-00993-8>
- Pedersen, S. M., Damslund, N., Kjær, T., & Olsen, K. R. (2025). Optimising test intervals for individuals with type 2 diabetes: A machine learning approach. *PLOS ONE*, 20(2), Article e0317722. <https://doi.org/10.1371/journal.pone.0317722>
- Pourdaryaei, A., Mohammadi, M., Mubarak, H., Abdellatif, A., Karimi, M., Gryazina, E., & Terzija, V. (2024). A new framework for electricity price forecasting via multi-head self-attention and CNN-based techniques in the competitive electricity market. *Expert Systems with Applications*, 235, Article 121207. <https://doi.org/10.1016/j.eswa.2023.121207>
- Rabi'atul'adawiah Shabli, Ul-Saufie, A. Z., & Ibrahim, N. (2025). Improving the prediction of the labour skill classification model in Malaysia using tree-based machine learning algorithms. *Malaysian Journal of Fundamental and Applied Sciences*, 21(5), 2771-2789. <https://doi.org/10.11113/mjfas.v21n5.4435>
- Rafrastara, F. A., Ghozi, W., Sani, R. R., Handoko, L. B., Abdussalam, A., Pramudya, E. R., & Abdollah, F. M. (2025). Integrating information gain and chi-square for enhanced malware detection performance. *Journal of Information and Communication Technology*, 24(1), 80-104. <https://doi.org/10.32890/jict2025.24.1.4>
- Sakyi-Yeboah, E., Agyemang, E. F., Agbenyeavu, V., Osei-Nkwantabisa, A., Kissi-Appiah, P., Moshood, L., Agbota, L., & Nortey, E. N. N. (2025). Heart disease prediction using ensemble tree algorithms: A supervised learning perspective. *Applied Computational Intelligence and Soft Computing*, 2025(1), Article 1989813. <https://doi.org/10.1155/acis/1989813>
- Salah-Ud-Din, M., S. S., B. T. L., & Al Ali, H. (2024). Exploratory data analysis and prediction of passenger satisfaction with airline services. In *2024 New Trends in Civil Aviation (NTCA)* (pp. 295-302). <https://doi.org/10.23919/NTCA60572.2024.10517814>
- Sani, J., Halane, S., Ahmed, A. M., & Ahmed, M. M. (2025). Application of machine learning algorithms and SHAP explanations to predict fertility preference among reproductive women in Somalia. *Scientific Reports*, 15(1), Article 26301. <https://doi.org/10.1038/s41598-025-04704-y>

- Sariman, S. R., Ab Jalil, H., & Marlisah, E. (2024). Prediction model of school dropout factors using classification techniques in Selangor. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, 9(6), Article e002867. <https://doi.org/10.47405/mjssh.v9i6.2867>
- Shaha, T. R., Begum, M., Uddin, J., Torres, V. Y., Iturriaga, J. A., Ashraf, I., & Samad, M. A. (2024). Feature group partitioning: An approach for depression severity prediction with class balancing using machine learning algorithms. *BMC Medical Research Methodology*, 24(1), Article 123. <https://doi.org/10.1186/s12874-024-02249-8>
- Shapley, L. S. (1953). *A value for n-person games*. Defence Technical Information Centre. <https://doi.org/10.1515/9781400881970-018>
- Sheokand, S., & Beniwal, S. (2025). Credit card fraud detection using machine learning and deep learning techniques: A review. *Pertanika Journal of Science and Technology*, 33(6). <https://doi.org/10.47836/pjst.33.6.11>
- Soni, K., Frew, R., & Kebede, B. (2026). Interpretable machine learning for origin classification of Brazilian soybeans: A random forest and XAI-based approach. *Journal of Food Composition and Analysis*, 150, Article 108896. <https://doi.org/10.1016/j.jfca.2026.108896>
- Syeda, D. Z., & Asghar, M. N. (2023). Enhancing malware classification: A comparison of filter and embedded feature selection with a novel dataset. In *Proceedings of the Cyber Research Conference* (pp. 1-8). <https://doi.org/10.1109/Cyber-RC159474.2023.10671445>
- Tsai, C.-F., Chen, K.-C., & Lin, W.-C. (2024). Feature selection and its combination with data over-sampling for multi-class imbalanced datasets. *Applied Soft Computing*, 153, Article 111267. <https://doi.org/10.1016/j.asoc.2024.111267>
- Ul Haq, E., Yong, Q., Yuan, Z., Jianjun, H., Ul Haq, R., & Qin, X. (2024). Accurate multiclassification and segmentation of gastric cancer based on a hybrid cascaded deep learning model with a vision transformer from endoscopic images. *Information Sciences*, 670, Article 120568. <https://doi.org/10.1016/j.ins.2024.120568>
- Unlu, A., & Subasi, A. (2025). Substance use prediction using artificial intelligence techniques. *Journal of Computational Social Science*, 8(1), Article 21. <https://doi.org/10.1007/s42001-024-00356-6>
- Visibelli, A., Finetti, R., Roncaglia, B., Poli, P., Spiga, O., & Santucci, A. (2025). Predicting therapy dropout in chronic pain management: A machine learning approach to cannabis treatment. *Frontiers in Artificial Intelligence*, 8, Article 1557894. <https://doi.org/10.3389/frai.2025.1557894>
- Wang, H., Liang, Q., Hancock, J. T., & Khoshgoftaar, T. M. (2024). Feature selection strategies: A comparative analysis of SHAP-value and importance-based methods. *Journal of Big Data*, 11(1), Article 44. <https://doi.org/10.1186/s40537-024-00905-w>
- Wang, P., Chen, T.-H., Chang, M.-Y., Hsia, H.-Y., Dai, M., Liu, Y., Hsu, Y.-L., Fu, F., & Zhao, Z. (2025). An explainable artificial intelligence framework for weaning outcomes prediction using features from electrical impedance tomography. *Computer Methods and Programs in Biomedicine*, 267, Article 108811. <https://doi.org/10.1016/j.cmpb.2025.108811>
- Wang, Z., He, C., Hu, Y., Luo, H., Li, C., Wu, X., Zhang, Y., Li, J., & Cai, J. (2024). A hybrid deep learning scheme for MRI-based preliminary multiclassification diagnosis of primary brain tumours. *Frontiers in Oncology*, 14, Article 1363756. <https://doi.org/10.3389/fonc.2024.1363756>

- Zahra, S. T., Imdad, S. K., Khan, S., Khalid, S., & Baig, N. A. (2025). Power transformer health index and life span assessment: A comprehensive review of conventional and machine learning-based approaches. *Engineering Applications of Artificial Intelligence*, 139, Article 109474. <https://doi.org/10.1016/j.engappai.2024.109474>
- Zeffora, J., & Shobarani, S. (2022). Optimising random forest classifier with Jenesis-index on an imbalanced dataset. *Indonesian Journal of Electrical Engineering and Computer Science*, 26(1), 505-511. <https://doi.org/10.11591/ijeecs.v26.i1.pp505-511>
- Zermane, A., Mohd Tohir, M. Z., Zermane, H., Baharudin, M. R., & Mohamed Yusoff, H. (2023). Predicting fatal fall from heights accidents using random forest classification machine learning model. *Safety Science*, 159, Article 106023. <https://doi.org/10.1016/j.ssci.2022.106023>
- Zhao, G., & Li, D. (2025). TreeEM: Tree-enhanced ensemble model combining with feature selection for cancer subtype classification and survival prediction. *Results in Applied Mathematics*, 27, Article 100605. <https://doi.org/10.1016/j.rinam.2025.100605>
- Zhao, X., Wu, D., & Wang, L. (2025). Wage effect and workers' skill mismatch: Paths and mechanisms. *SAGE Open*, 15(1), Article 21582440251318540. <https://doi.org/10.1177/21582440251318540>
- Zhou, S., & Mentch, L. (2023). Trees, forests, chickens, and eggs: When and why to prune trees in a random forest. *Statistical Analysis and Data Mining*, 16(1), 45-64. <https://doi.org/10.1002/sam.11594>