Journal of Information Systems Engineering and Management
2025, 10(29s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Enhancing Teacher Placement in Workforce Management using the Staff Placement Model and J48 Classification

Shamsul Saniron[1], Zulaiha Ali Othman[2], Abdul Razak Hamdan[3], Wan Nur Syamilah Wan Ali[4]
[1, 2, 3, 4] University of National Malaysia

ARTICLE INFO

Received: 28 Dec 2024

Revised: 18 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

Introduction: Effective teacher placement plays a crucial role in ensuring workforce efficiency, improving the quality of education, and enhancing institutional excellence. However, the Ministry of Education Malaysia (MOE) continues to face challenges in aligning teachers' expertise with school needs. The reliance on administrative considerations through the Teacher Placement Committee (TPC) without predictive capabilities has led to inefficiencies and suboptimal assignments.

Objectives: To develop a systematic, data-driven teacher placement model using classification techniques to enhance decision-making efficiency and to evaluate its effectiveness by comparing its placement decisions with those made by the Teacher Placement Committee (TPC).

Methods: This study employs a methodology that includes a literature review, model development, data collection, and model validation. Data mining techniques were used to analyze teacher placement patterns. The proposed model, adapted from the Staff Placement Model (SPM), integrates factors such as subject specialization, workforce demand, and institutional needs. An experiment was conducted using 5,526 teacher placement applications from the year 2022, considering four key components: demographics, environment, staffing requirements, and human factors

Results: The results indicate that the J48 technique achieved the highest accuracy (96.30%) compared to Decision Tree, Naïve Bayes, Random Forest, and K-Nearest Neighbors. The model's effectiveness was further evaluated using 3,096 new application records. The results showed that the proposed model approved 1,554 placements compared to 1,5421 in the actual decisions. The t-test results also showed a high p-value ($p > 0.05$), proving that the model's decisions were comparable to those made by the TPC with no significant difference.

Conclusions: In conclusion, the developed staff placement model can be effectively applied to develop a teacher placement model (TPM) at MOE. This TPM model has the potential to serve as an essential decision-support tool for MOE in improving the efficiency of teacher placement in the future.

Keywords: Teacher placement, classification techniques, data mining, J48

INTRODUCTION

The role of teachers as a key factor of human resource management (HRM) in the education field must be carefully placed because it relates directly with job performance, learning outcomes, and institution effectiveness. Proper teacher allocation ensures that teachers are placed in jobs that are aligned with their field of study, expertise, and what the school needs. However, current placement systems are frequently inefficient, resulting in personnel imbalances, mismatch of teaching professionals, and poor education quality [4].

Use subject-based teachers as Malaysia's MOE would want to give teachers based on the subjects they know best after their pre-service training was done [6]. Despite all these efforts, mismatches still occur, making teachers unsatisfied and school management less effective [8]. This identifies the need for a fair, structured data-driven process for teacher allocation that would help in improving teacher deployment.

The previous method of teacher allocation has numerous issues, particularly when the Staff Placement Committee (SPC) is required to rummage through voluminous files to access details. It is difficult to comprehend, time-consuming and prone to errors, which could result in the recruitment of teachers not matching the school's needs [1].

The placement of teachers is based on something like administrative tests and, even worse, people's own opinions, with the assistance of systems such as HRMIS [9], e-Operasi [3], e-Gtukar [2]. But these platforms are primarily just used to record things and not to make intelligent decisions. Placement choices are reactive, inconsistent, and inefficient because predictive analytics are not used. This complicates things for teachers who are trying to align their skills to the needs of the school.

Data mining is used by businesses to sift through large solutions of data to identify trends that allow them to make more informed decisions. Incorporating data sciences in teacher hiring is common in health care and marketing [4], but not in teacher hiring. The system can apply data mining to determine the best way to assign teachers, taking into account past success and the needs of schools and the region from which they are drawn [20] There still are teacher-specialty mismatches, which leads to ineffective teaching, student unhappiness and regional issues [11].

At this point in time, placement systems do not predict future outcomes, leaving them dependent on manual decisions driven by short-term demands. This prevent functioning is time-consuming, leading to assignments that are not only less consistent [15]. But employing data-driven insights, data mining can optimize the placement of teachers, reduce mismatches, and ensure balanced staff.

OBJECTIVES

This study develops a data-driven teacher placement model (TPM) using classification techniques to improve decision-making. Its effectiveness is evaluated by comparing its placements with actual Teacher Placement Committee (TPC) decisions.

TPM enhances accuracy through structured data integration and classification techniques. Key data from MOE systems (e-Operasi, e-GTukar, HRMIS, and EMIS) refine decision-making. An experimental approach identifies the best classification model, validated through historical data testing.

By leveraging data mining and predictive modeling, TPM improves efficiency, fairness, and accuracy, minimizing mismatches and ensuring equitable teacher distribution.

LITERATURE REVIEW

2.1 Placement Teachers

This practice is integral to affirming the efficiency of the workforce, the quality of education, and ensuring excellence at the institution. The Ministry of Education Malaysia (MOE) has been facing challenges in aligning teachers' expertise with schools' demands. With predictive capacity, we can move away from the administrative TPC-driven placements that currently lead to inefficiencies and poor placements.

The TPM is developed through a systematic and structured manner to ensure the optimal placement of teachers based on their subject area profile, workforce demand and institutional requirements. The model employs a data-driven methodology within a decision-making framework to improve accuracy, fairness, and efficiency in the placement of Teachers [19].

2.2 Data Mining in Theachears Placement

Data Mining is the method of discovering patterns between data and can also be termed as knowledge discovery from (kDD). Data mining helps maximize teacher placement through data-driven approaches that analyze historical data to determine significant aspects impacting teacher allocation. Data mining can study the qualifications of assigned teachers, contracts, specialty knowledge, and school needs to help develop a distribution system of human resources, reduce mismatches in subject specialization [9].

The use of machine learning-based data mining techniques in teacher placement is informative as well and includes classification, forecasting, cling, and anomaly detection [28]. For example, classification predicts optimal assignments between teachers and schools, forecasting predicts potential shortages or oversupply in schools, and clustering groups schools together that have similar demand and geographic considerations to ensure teachers are distributed appropriately.

Data mining has been previously employed analyzing human resources for workforce allocation, however mainly on general staff and not teachers. Although predictive analytics is widely applied to HRM, there has been little research on teacher placement with knowledge discovery methods [16]. Teacher placement is usually across all ten states and at twenty-two states is ineffective without a disciplined, data-driven approach that can lead to supply-demand mismatches and imbalance in workforce. We aim to address this issue through data mining to enhance the accuracy and fairness of teacher placement. Using data mining as a solutions approach to mismatched teacher assignment. Patterns recognition allows us to know what contributed to bad placements, and how to enhance the placement strategy based on the data [11]. As the data are used to solve the problem of teacher allocation, a data-driven model can accelerate the allocation process, minimize time spent on manual decisions, and ensure fewer errors are made.

2.2.1 Data Mining Techniques in Applications

Data mining (DM) techniques are predominantly used for forecasting and are becoming increasingly prominent in the analysis of practical big data. A large number of studies have been performed to classify DM techniques. The J48 method is suitable for outcomes where classification is necessary, easy to use, yet also interpretable, even when dealing with large datasets. It produces classification rules in the form of tree units or rule sets [5]. Studies have proven Decision Tree classification particularly J48 improves prediction power. J48 is famous and popular as it creates rules that don't require domain expertise.

J48 is frequently used for classification and prediction as reported in previous studies. Although the output of decision trees is quite simple, it is also easy to use and interpret; thus, it is often used to produce decision rules [21]. This is because J48 is an algorithm based on popular classification methods that utilize multiple information sources.

1. Select the attribute that has the highest multiple information
2. S contains s_i tuples for class C_i for $i = 1, 2, \dots, m$
3. A measure of information or expected information is needed to classify an arbitrary tuple as Equation (1):
$$I\ (s_1, s_2, \dots, s_m) = -\sum_{i=1}^{m} \frac{s_i}{s} log2 \frac{s_i}{s} \quad (1)$$
4. Entropy for attribute A with values $a_1, a_2, \dots, a_v$ as Equation (2):
$$E\ (A) = \sum_{j=1}^{v} \frac{s_{1j} + \dots + s_{mj}}{s} I\ (s_{1j}, \dots, s_{mj}) \quad (2)$$
5. Information acquisition means how many multiples will be obtained by pruning attribute A as Equation (3):
$$Gain\ (A) = I\left(s_1, s_2, \dots, s_{mj}\right) - E\ (A) \quad (3)$$

The J48 technique leverages the concept of information gain to produce easily understandable rules in the form of a decision tree. These classification rules represent the implicit knowledge within the database and are then used to classify and predict outcomes for new or unseen data, beyond the training set.

2.2.2 Data Mining in HRM

DM analysis here is one of the best methods to analyze data from databases which can help decide on further planning. The use of DM methods for staff placement in HRM is much less established. Research on DM approaches employed in HRM has received relatively less attention among other fields. This study therefore addresses DM methods for staff placement. Relevant HRM applications in this scope have been defined in Table 1.

TABLE 1: Data Mining Techniques in HRM

Technique	HRM Application
Decision Tree	Evaluation and Analysis of Human Resource Management Models and Talent Selection Factors [26], Analysis of Human Resource Management Models and Selection Criteria [13].
Association Rule Mining	Enhanced Association Rules Algorithm and Cloud Service System for HRM [24], Optimization of Human Resource Allocation [23].

Technique	HRM Application
Rough Set Theory	Automated Academic Staff Performance Evaluation System Using Rough Sets Theory [18], Key Factors for Success in New Service Development [17].
Fuzzy DM	Evaluation of Tourism Human Resource Management Performance [22], Early Warning Management of Enterprise Human Resource Crises [23].
Artificial Neural Network	Analysis and Simulation of an Early Warning Model for HRM Risks [25], Evaluation and Image-Based Analysis of Enterprise HRM [27].

2.3 Staff Placement Model in Human Resource Management

Staff placement is an important HRM function, to ensure employees are assigned to positions that match their skills, qualifications and organizational needs. A functional placement system strategically utilizes a workforce, fosters job satisfaction, and ultimately benefits the organization as a whole. However, traditional manual methods of placement often suffer from issues like inconsistent decision-making, absence of standard criteria, and inefficient sourcing of work force [20]. To address these challenges, structured Strategic Personnel Management (SPM) approaches have emerged, leveraging data-driven applications, frameworks, and solutions to enhance placement accuracy, reliability, and overall effectiveness. The Personnel Placement Model is illustrated in Figure 1

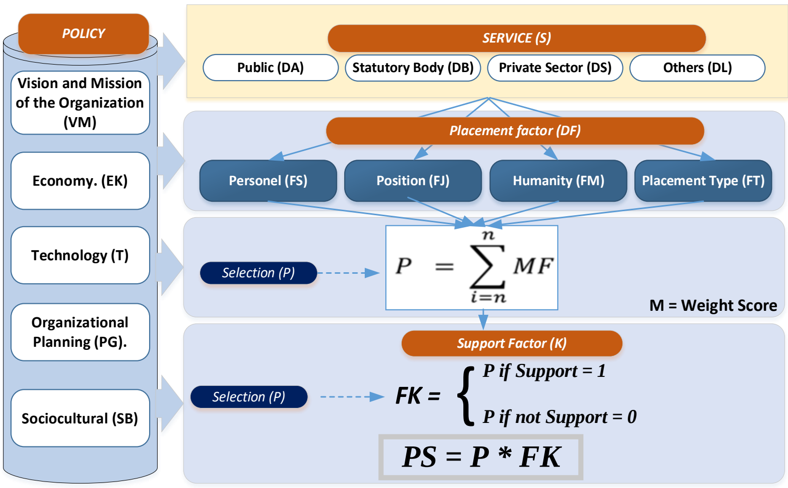

FIGURE 1 : Staff Placement Model (SPM)

SPM is a systematic method of allocating workforce taskingsin large organizations. It leverages human expertise-driven decision criteria to ensure staff placements address organizational-level mandates and workforce distribution needs. SPM uses data mining techniques to make Placement Decisions that are more systematic, scalable and arelinked to long-term Workforce Planning.

The SPM holds key factors impacting allocation decisions. The first is policy factors, which setorganizational strategies and guidance. These include workforce planning, financial constraints, digital evolution as well as socio-cultural factors. One solution to high turnover is cross-departmental training that allows for fluid placement so that the needs of the organization can be met.

TPM is a tailored version of SPM designed for teacher placement. TPM attributes are derived from attributes in SPM. Table 2 shows the corresponding attributes available in the teacher application database.

TABLE 2 Corresponding attributes found in the teacher application database.

Factor Code	Factor	Total Attributes
FS	Personel/ Demographic	17
FJ	Position/ : Staffing Requirement	4
FT	Placement Type	3

FM	Humanity	14

METHODS

3.1 Research Methodology Framework

This study follows a structured methodology divided into four main phases to develop a TPM using data mining techniques. The TPM is adapted from the existing SPM, incorporating modifications to cater specifically to teacher placement within the MOE. The methodology ensures that the model is data-driven, optimized for structured decision-making, and applicable in real-world teacher allocation processes. Figure x shows the overall research framework.

Phase 1: Preliminary Study & Literature Review
This phase includes a literature review on data mining applications, HRM, and SPM to ensure TPM development aligns with the education sector.

Phase 2: Teachers Placement Model (TPM) in MOE
This phase focuses on designing and developing TPM based on insights from the literature review. It begins by identifying key factors influencing teacher placement. To ensure practical relevance, expert interviews were conducted to select suitable attributes for the TPM.

Phase 3: Data Collection & Model Implementation Using Data Mining (Case Study)
This phase involves data collection and preparation for developing and validating TPM Run the classification algorithms J48, Decision Tree, Naïve Bayes, Random Forest, and Kstar using Weka 3.8.6

Phase 4: Evaluation of the TPM
The TPM was validated using second face 2022 teacher placement data to test classification algorithms. The best technique was selected based on accuracy in predicting placements. Validation compared actual and predicted assignments to measure efficiency in reducing mismatches.

3.2 Teachers Placement Model (TPM) in MOE

This phase is analyzes teacher placement data from 2022, focusing on key attributes used in school management. The most relevant attributes, derived from SPM, guide placement decisions.

To validate these attributes, 10 experts were interviewed, providing feedback and recommendations. The selected attributes were matched with database records, followed by data analysis. Data mining techniques were then applied to create the best placement model and rules. Finally, a comparison of user satisfaction between the existing system and TPM was conducted. Figure 2 shows the TPM development phases.

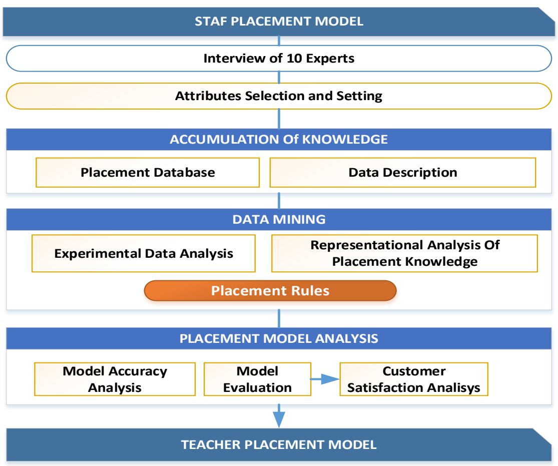

FIGURE 2: TPM Development Phases

3.2.1 Placement Attributes

To validate TPM in order to keep it aligned and up to date, a formal expert validation process was held. This study adapted those placement attributes in TPM from the previously developed SPM to be consistent with existing frameworks for workforce allocation.

These tables are presented in more detail in Tables 3, 4, 5 and 6 for each factor selected. These tables contain the validated teacher placement attributes, their categorization, and importance in TPM. This logical approach ensures that teachers are placed in Malaysian schools fairly, efficiently, and accurately.

TABLE 3: Personel/ Demographic Attributes

Id	Attribute Name	Details
FS1	Gender	Male / female
FS2	Marital status	0: No information, 1: Not married, 2: married, 3: widowed, 4: Widowed, 5: widowee 9: other
FS3	Position	According to the number of posts: 54 types of posts
FS4	Original option	Option code
FS5	Dominant option	Dominant Option Codes: 76 types of Options
FS6	Contractual bond	Yes / no
FS7	Period of service in the state	Date of placement in the state
FS8	Period of service in the district	Date of settlement in the district
FS9	Teaching subject Experiencee	Subject code
FS10	Long tenure	Month
FS11	Long apart	Year / month
FS12	Country of origin	State code
FS13	Original district education office (PPD)	District name / code applied

TABLE 4: Position/ Staffing Requirement Attributes

ID	Attribute Name	Details
FJ1	Type of position	1: Administrative position, 2: non-administrative position
FJ2	Staffing requirements according to staffing norms	Information on the status of staffing requirements
FJ3	Staffing requirements according to option	Information on the status of staffing requirements
FJ4	Staffing requirements according to position	Information on the status of staffing requirements

TABLE 5: Placement Type Attributes

ID	Attribute Name	Details
FT1	Application state	State name / code applied
FT2	Type of school	Primary, secondary, full boarding, religious, vocational, art, sports
FT3	Application PPD	PPD code

TABLE 6: Description of Humanitarian Attributes

ID	Attribute Name	Details
CASE_1	Chronic pain case	1: Yes, 0: No (weightage: 25)
	Chronic pain case: self	1: Yes, 0: No (weightage: 50)

ID	Attribute Name	Details
	Chronic pain case: spouse	1: Yes, 0: No (weightage: 40)
	Chronic pain case: Children	1: Yes, 0: No (weightage: 40)
CASE 2	Personal safety threats	1: Yes, 0: No (weightage: 25)
	Personal safety threats: death threats	1: Yes, 0: No (weightage: 50)
	Personal safety threats: injury threats	1: Yes, 0: No (weightage: 30)
CASE_3	Follow spouse	1: Yes, 0: No (weightage: 20)
	Follow spouse: instruction	1: Yes, 0: No (weightage: 40)
	Follow spouse: according to spouse in the state / PPD / school applied for	1: Yes, 0: No (weightage: 30)
	Follow spouse: joint application	1: Yes, 0: No (weightage: 10)
CASE_4	Term of service	1: Yes, 0: No (weightage: 20)
	Period of service: less than 4 years in state / PPD / school	1: Yes, 0: No (weightage: 5)
	Period of service: more than 4 years in the state / PPD / school	1: Yes, 0: No (weightage: 10)

3.2.2 Interview

The interview process was conducted with 10 experts. The structured interviews were conducted face-to-face, where experts were provided with a predefined list of attributes based on SPM and asked to evaluate and select the most relevant attributes for teacher placement. Table 7 shows the list of experts by position and organization. Each expert reviewed and marked attributes based on their importance in teacher placement, ensuring a broad consensus on essential factors for decision-making. The final selection process retained key placement attributes related to personnel factors, job requirements, policy considerations, and humanitarian concerns.

TABLE 7: List of Experts

Expt	Position	Organization
1	Deputy Director of the School Management Sector	Daily School Management Department
2	Chief Assistant Director Projection Unit Education Macro Planning Sector	MOE
3	Assistant Director Projection Unit Education Macro Planning Sector	MOE
4	Member of the Malaysian Service Commission	Malaysia Service Commission
5	Head of Unit Human Resource Management	IPG
6	Chief Assistant Secretary Service Section	JPA
7	Deputy Director of the Division Service Branch	MOE
8	Chief Assistant Secretary Teacher Recruitment Unit	MOE
9	Assistant Director	State Education Department
10	Chief Assistant Director	Daily School Management Department

Table 8 provides the validation results indicating both the consistency and the significance of certain characteristics across teacher placement decisions, confirming, for example, that factors such as job category, employment status, marital status, teaching options, and geographical preferences were critical. In addition to work experience, subject area specialization, and length of service at the current school, which are very practical aspects that affect the teaching quality and are for planning the workforce, the experts also highlighted the importance of the instrument when doing any kind of assessment. Also, humanitarian grounds, especially on health issues were among the major determinants in granting teacher transfers.

The approved characteristics were integrated into the definitive TPM to facilitate a data-based, rational and transparent teacher distribution process. While placement would be substantially based on these characteristics, the details of the respective placements would follow a structured matching process using workforce data and strategic planning needs.

TABLE 8: List of Selected Attributes for Teacher Placement Model Process

Component	ID	Attribute Name	e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	Status
Person-el/ Demographic	FS1	Gender	-	/	/	-	-	/	/	-	-	-	
	FS2	Marital status	/	/	/	/	/	/	/	/	/	/	Y
	FS3	Position	/	/	/	/	/	/	/	/	/	/	Y
	FS4	Original option	-	/	-	-	/	-	-	/	-	-	-
	FS5	Dominant option	/	/	/	/	/	/	/	-	/	/	Y
	FS6	Contractual bond	-	-	-	-	-	-	-	-	-	-	-
	FS7	Period of service in the state	/	/	/	/	/	/	/	/	/	/	Y
	FS8	Period of service in the district	/	/	/	/	/	/	/	/	/	/	Y
	FS9	Teaching subject Experiencee	-	-	-	-	-	-	-	-	-	-	
	FS10	Long tenure	/	/	/	/	/	/	/	/	/	/	Y
	FS11	Long apart	/	/	-	-	-	-	/	-	-	-	-
	FS12	Country of origin	/	-	-	-	-	-	-	-	-	-	-
	FS13	Original district education office (PPD)	/	-	-	-	/	-	-	-	-	-	-
Position / : Staffing Requirement	FJ1	Type of position	/	/	/	/	/	/	/	/	/	/	Y
	FJ2	Staffing requirements according to staffing norms	/	/	/	/	/	/	/	/	/	/	Y
	FJ3	Staffing requirements according to option	/	/	/	/	/	/	/	/	/	/	Y
	FJ4	Staffing requirements according to position	/	/	/	/	/	/	/	/	/	/	Y
Placement Type	FT1	Application state	/	/	/	/	/	/	/	/	/	/	Y
	FT2	Type of school	/	/	/	/	-	/	/	-	/	/	Y
	FT3	Application PPD	/	/	/	/	/	/	/	/	-	/	Y

e = Expert Y = Select

Based on the interviews conducted, the experts selected 13 key attributes to be used in the placement selection process.

3.2.3 Accumulation of Knowledge

The process of teacher placement, underpinned by the MOE's accumulation of knowledge, is key to ensuring a systematic, fair and efficient distribution of educators throughout schools. With real-time data analysis, MOE can distribute the teachers equally around the country; and every school has enough teachers, and all the teachers are qualified in that subject.

Knowledge accumulation serves one of its main purposes in the area of teacher placement and teacher movement by preventing the imbalance in the distribution of teachers in urban and rural schools. MOE uses historical placement data to find schools with teacher shortages, and assigns new teachers accordingly. For example, if there is data showing that a lot of rural schools have no or few Science teachers, camp MOE to take precautionary action and increase Science teachers to the area as well as raise the quality of service.

A comprehensive knowledge base will also enable more efficient handling of teachers placement. However, with a centralized and systematic database, a much more objective and accurate placement decision is possible, minimizing mismatched assignments that can potentially cause a high turnover of teachers. For instance, data analytics and artificial intelligence (AI) can be utilized by the MOE to identify the best school for each teacher based on experience, subject expertise, and geographical desirability.

In addition,piled-up knowledge rises the satisfaction and motivation of the teacher. MOE can optimize this process by taking into account location preferences (ie distance from home) as well as existing infrastructure for teaching & learning, to train teachers in areas that are relevant for them. That makes it less stressful when placed in remote location for they reduce the spread stress and feel happy with their job making them better teachers.

a) Data Description refers to the collection, classification, and analysis of information related to teachers and school requirements. This data is used to understand trends, patterns, and challenges in the teacher placement process.

Key Components of Data Description:

1. Teacher Information
   - Name, age, gender, and service status.
   - Academic qualifications and areas of specialization.
   - Teaching Experiencee and performance records.

2. Placement Information
   - School location (urban/rural).
   - Number of existing teachers by subject.
   - Demand for teachers in each school.

3. Job Demand & Vacancies
   - Subjects facing a shortage of teachers.
   - Teachers requesting transfers or resignations.

4. Trend Analysis & Future Projections
   - Data on teachers retiring within the next 5-10 years.
   - Past teacher placement patterns and their effectiveness.

b) Data preprocessing is a crucial step to enhance data quality before analysis. Below are the main techniques used:

1. Handling Outliers
   Outliers are extreme values that can impact analysis results. The number of outliers detected using the Interquartile Range (IQR) method is as follows. Table 9 shows the Number of Outliers

TABLE 9: Number of Outliers

Column	**Number of Outliers**
Service_Duration	290
Humanity_Weight	163
Experiencee	0

Handling method:

- "Service_Duration" & " Humanity_Weight " → **Winsorization** (replacing extreme values with upper and lower limits).
- "Experiencee" → **No action needed**, as no outliers were detected.

The IQR method is used to identify and remove outliers based on the interquartile range.

Formula:

$$IQR = Q3 - Q1$$
$$\text{Lower Bound} = Q1 - 1.5 \times IQR$$
$$\text{Upper Bound} = Q3 + 1.5 \times IQR$$

Where:

- Q1 (First Quartile) = the value below which 25% of the data falls
- Q3 (Third Quartile) = the value below which 75% of the data falls
- IQR = the difference between Q3 and Q1

Processing options:

- Remove values outside the Lower Bound and Upper Bound.
- Replace outliers with the Q1 or Q3 value.

2. Normalization
   Normalization helps ensure that data is on a consistent scale, especially for machine earning.

   Handling method:
   - Min-Max Scaling is used to scale data into the range **[0,1]**.
   - It is applied on numeric columns such as " Service_Duration ", "Humanity_Weight", dan " Experiencee ".

Formula used: (Min-Max Scaling (Normalization))

$$Xnorm = \frac{X - Xmin}{Xmax - Xmin}$$

- Xnorm = Normalized value
- X = Original value
- Xmin = Minimum value in the dataset
- Xmax = Maximum value in the dataset

Function: Scales data into the range [0,1] or another specified range. **Advantages: Useful** when data does not follow a normal distribution. Disadvantages: Highly sensitive to outliers because it depends on the maximum and minimum values.

c) Placement Database is a database system that digitally stores and manages all information related to teacher placements. This system helps the MOE make placement decisions based on accurate and up-to-date data. Figure 3 shows the ERD diagram for the Placement Database System.

Key Functions of the Placement Database:

1. Teacher Registration & Data Storage
   - Stores profiles of teachers, including their qualifications and Experiencee.
2. School Needs Mapping
   - Identifies schools facing teacher shortages based on subject requirements.
3. Placement Recommendation System
   - Uses smart algorithms to suggest the most suitable placement locations.
4. Online Teacher Transfer Application
   - Allows teachers to apply for transfers online based on specific criteria.

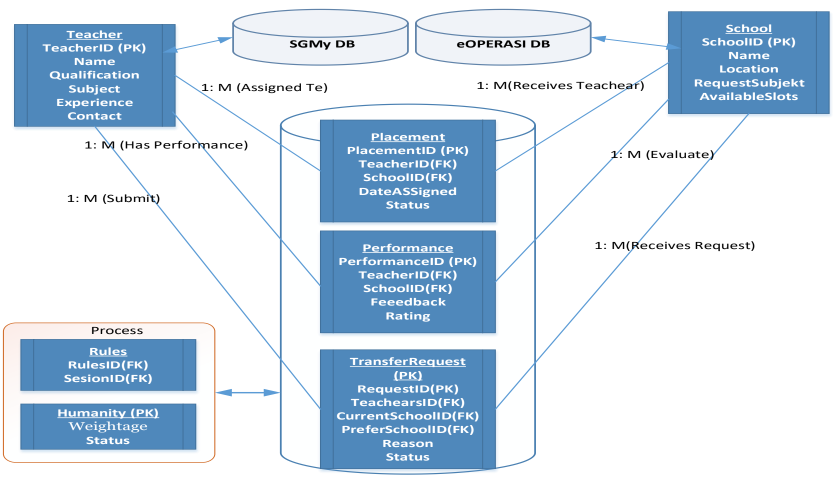

FIGURE 3: Diagram for the Placement Database System.

5. Data Integration
   - Dataset Combination: Merging data from multiple sources, including the eOperasi, eGTUKAR, and SGMy databases.
   - Eliminating Data Inconsistencies: Aligning data from various sources to ensure consistency.
   - Manual File Combination: Utilizing Microsoft Excel, Access, and SQL to manually combine files.
   - Integration of Teacher Data: Merging teacher data from various states into a unified file.
   - We are also standardizing column names and formats by changing column names, data formats, and data types to make sure they are uniform.
   - Cleaning and Adjusting of Data: Cleaning the data and making it standardized.
   - Data Storage: Storing the data into a storage system or a database.

   An important aspect of ensuring that Malaysia's education system runs effectively is through the form of knowledge accumulation, in the form of teacher placement. The role of MOE in collecting and consuming this data ensures that teachers can be distributed wisely, specialization match with the subjects taught, and that the overall efficiency of teacher placement is improved. It is a win-win situation for the teachers, students and the education sector as a whole leading to a more organized, data-driven and value-added system in teaching and learning in days to come.

3.2.4 Data mining

We perform data mining for TPM in an organized manner, consisting of data acquisition and preprocessing, classification, and validation. The primary data we used was collected from eGTUKAR which is a system used to manage teacher transfer applications that come under the MOE. This step aims to find usable patterns in the data to optimize teacher placement.

In the first phase, data collection was conducted by retrieving teacher application data from the eGTUKAR system. This data set consisted of teacher demographics, including placements, work experience, subject specialty, and humanitarian factors influencing transfers, for example — medical issues or requests to transfer, due to spouse relocation. Because teacher transfers happen biannually, the data was sufficiently deep to capture teacher movement phenomena. In order to process it more efficiently, we removed redundant data and divided the dataset into 90% training data and 10% test data for model evaluation. For model robustness, a 10-fold cross-validation was used. The following classification algorithms were tested: J48, Decision Tree, Naïve Bayes, Random Forest, and Kstar. These models were used to ascertain the best technique to predict teacher placements from historical data.The analysis phase of the TPM focuses on evaluating the accuracy, classification techniques, and predictive performance of the model. This phase ensures that the selected classification algorithms can make reliable and precise teacher placement predictions while optimizing workforce allocation.

In this section, we present the placement data used in this case study. Records of 5,526 inter-state transfer and placement application records collected during the year 2022 were analyzed. These records are applications for the teaching staff group to be transferred to a new placement area. Table 10 displayed 13 selectable attributes by subjective experts and 4 superficial attributes from the humanity factor in Table 6. The target class indicates the output.

TABLE 10: Attributes Type, Number & Description

Component	ID	Attribute Name
Personel/ Demographic	FS2	Marital status
	FS3	Position
	FS5	Dominant option
	FS7	Period of service in the state
	FS8	Period of service in the district
	FS10	Long tenure
Position/ : Staffing Requirement	FJ1	Type of position
	FJ2	Staffing requirements according to staffing norms
	FJ3	Staffing requirements according to option

Component	ID	Attribute Name
	FJ4	Staffing requirements according to position
Placement Type	FT1	Application state
	FT2	Type of school
	FT3	Application PPD
Target		Placement Target (Successful @ Unsuccessful)

The experimental structure involved preparing the dataset by selecting relevant placement attributes, formatting the data for processing, and applying classification techniques to generate placement rules.

By utilizing advanced classification techniques, the TPM optimizes teacher placement decisions, minimizes human error, and enhances efficiency in Malaysia's teacher allocation system. The structured data-driven approach ensures that placement decisions are transparent, objective, and aligned with workforce distribution needs.

RESULTS

Classification was performed using the 10-fold cross-validation method. Accuracy was calculated by averaging the accuracy of classification models across 10 random datasets (R). The models were developed based on the number of training data (L) and testing data (U). This calculation method was consistently applied across all experiments to evaluate model accuracy. At this stage, experiments were conducted using Weka software version 3.8.6 (http://www.cs.waikato.ac.nz/ml/weka/). Five classification algorithms were applied to determine the highest accuracy result. Accuracy evaluation was performed using the cross-validation method. The cross-validation method is a statistical algorithm that divides the data into two segments: one used for training the model and the other for validating the model [12]. Table 11 shows the cross-validation method generated using the J48 technique.

TABLE 11 : Classification Accuracy of J48

Bil	% Split	*Correctly Classified Instances*	*Incorrectly Classified Instances*
1	**10 90**	**94.66 %**	**5.34 %**
2	20 80	93.56 %	6.44 %
3	30 70	93.56 %	6.44 %
4	40 60	93.56 %	6.44 %
5	50 50	92.82 %	7.18 %
6	60 40	94.17 %	5.83 %
7	70 30	93.93 %	6.07 %
8	80 90	94.29 %	5.71 %
9	90 10	93.44 %	6.56 %

The same process was applied to testing for each algorithm J48, Decision Tree, Naïve Bayes, Random Forest, and Kstar

The analysis of model accuracy from this experiment is summarized in table 12. The results for the full-attribute experiment indicate that the J48 algorithm achieved relatively higher accuracy compared to other classification algorithms across all three placement datasets used. The accuracy analysis for the full-attribute model also reveals that the number of data points impacts model accuracy. This is evident when comparing follow-up data to initial data, except for the J48 classification technique, which showed a slight decrease in accuracy. Additionally, the accuracy of models containing outliers was found to decrease slightly, highlighting the impact of outliers on model performance.

TABLE 12: Model Accuracy for Full Attributes

Algorithm	% Split		*Correctly Classified Instances*	*Incorrectly Classified Instances*
J48	10	90	94.66 %	5.34 %
K-Star	90	10	93.45 %	6.55 %
RF	30	70	93.87 %	6.13 %
Decision Table	20	80	92.54 %	7.46 %
NaiveBayes	10	90	93.89 %	6.11 %

The analysis aims to identify the number of rules derived from each dataset and determine the key attributes within them. These findings assist in evaluating and selecting the most relevant rules produced by the classifier. Additionally, the study assesses the rules using evaluation data, distinct from the training data, to determine the model's accuracy for new data. The J48 decision tree classification technique generates insights in the form of decision trees. Figure 4 and Figure 5 shows an example of a Decision Tree.

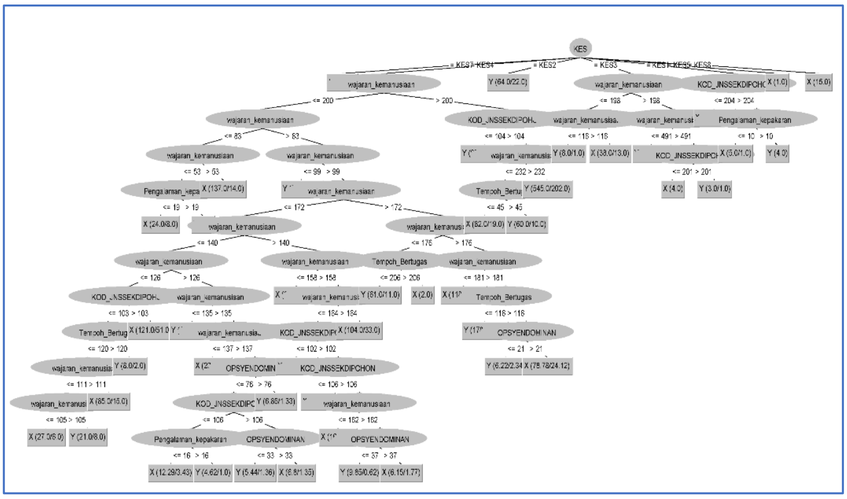

FIGURE 4: Example of a Decision Tree.

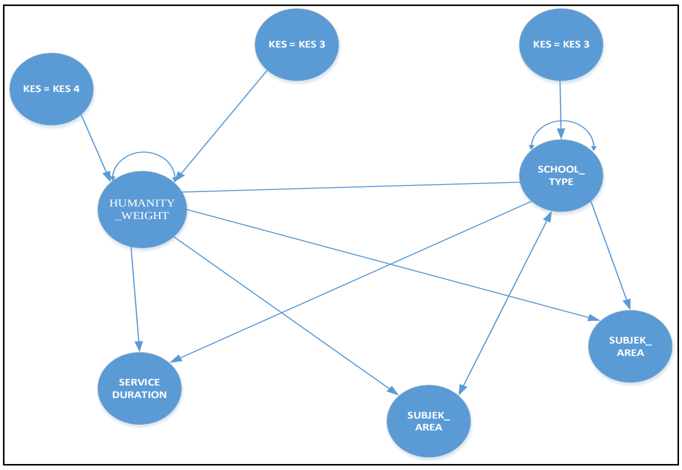

FIGURE 5: Decision Tree Visualization.

The decision tree consists of a set of **if-then** rules that classify data based on different attribute conditions. Here's how the rules can be interpreted. Table 14 show a decision rules for classification

TABLE 14 : Decision rules for classification

Condition	Decision (X/Y)
CASE = CASE7	Y
CASE = CASE2	Y
CASE = CASE5	X
CASE = CASE8	X
CASE = CASE1 AND SCHOOL_TYPE ≤ 204	X
CASE = CASE1 AND SCHOOL_TYPE > 204 AND Expertise_Experience≤ 10	X
CASE = CASE1 AND SCHOOL_TYPE > 204 AND Expertise_Experience> 10	Y
CASE = CASE3 AND Humanity_Weighting≤ 116	Y
CASE = CASE3 AND Humanity_Weighting> 116	X
CASE = CASE4 AND Humanity_Weighting≤ 53 AND Expertise_Experience≤ 19	X
CASE = CASE4 AND Humanity_Weighting≤ 53 AND Expertise_Experience> 19	Y
CASE = CASE4 AND Humanity_Weighting> 53 AND Humanity_Weighting≤ 83	X
CASE = CASE4 AND Humanity_Weighting> 83 AND Humanity_Weighting≤ 99	Y
CASE = CASE4 AND Humanity_Weighting> 99 AND Humanity_Weighting≤ 126 AND Duration of Service ≤ 120 AND Humanity_Weighting≤ 111	X
CASE = CASE4 AND Humanity_Weighting> 126 AND Humanity_Weighting≤ 135	Y
CASE = CASE4 AND Humanity_Weighting> 135 AND Humanity_Weighting≤ 137	X
CASE = CASE4 AND Humanity_Weighting> 137 AND SUBJEK_AREA ≤ 76 AND SCHOOL_TYPE ≤ 106 AND Expertise_Experience≤ 16	X
CASE = CASE4 AND Humanity_Weighting> 137 AND SUBJEK_AREA ≤ 76 AND SCHOOL_TYPE ≤ 106 AND Expertise_Experience> 16	Y
CASE = CASE4 AND Humanity_Weighting> 137 AND SUBJEK_AREA > 76	Y
CASE = CASE4 AND Humanity_Weighting> 140 AND Humanity_Weighting≤ 158	X
CASE = CASE4 AND Humanity_Weighting> 164	X

The decisions generated by this model depend on several key factors, primarily CASE category, Humanity_Weighting, Expertise_Experiencee, Duration of Service, and other factors such as SCHOOL_TYPE and SUBJEK_AREA.

Based on the extracted rules, CASE7 and CASE2 are automatically classified as Y (Yes) without the need to consider other factors. In contrast, categories like CASE5 and CASE8 are more likely to receive an X (No) decision. This indicates that certain CASE categories have an inherent advantage in the decision-making process.

Additionally, Humanity_Weighting plays a significant role in determining the final decision. If its value is low (e.g., ≤ 53), the applicant is more likely to receive an X (No) decision. However, as this value increases, the likelihood of obtaining a Y (Yes) decision also rises, especially if accompanied by higher Expertise Experiencee. In some cases, if the applicant's expertise Experiencee is low (e.g., ≤ 19 years), the decision may shift from Y to X, highlighting the importance of Experiencee in the evaluation process.

There are also critical thresholds in the decision-making process. For example, when Humanity_Weighting = 137, the final decision depends on SUBJEK_AREA and SCHOOL_TYPE. If these values fall within a specific range, the decision may shift from X to Y, showing that multiple factors interact to determine the final outcome.

Overall, the generated rules indicate that decisions are not based on a single factor but rely on a combination of several key variables. In some cases, applicants with higher values in Humanity Weighting and Expertise Experiencee have a greater chance of receiving a Y (Yes) decision, whereas applicants with lower values in these factors are more likely to receive X (No). The extracted rules also reveal consistent patterns that can be used to accurately predict outcomes based on the given data.

4.1 Integrated TPM in KPM

The staff placement model at KPM is based on the SPM model. The proposed TPM model integrates available KPM data sources, which are processed using intelligent techniques to uncover hidden insights, improve decision-making, and maximize the value of existing data resources. This study focuses on the attributes used in school management placement processes within KPM.

Figure 6 illustrates the TPM model implementation at KPM, adapted from the previously discussed SPM model. This diagram illustrates a data-driven teacher placement system. Information from the user databases (eGTUKAR, eOPERASI, SGMy, and transfer database) is channeled into the placement database for processing. The process includes placement rules, data mining, knowledge accumulation, and attribute selection before the final decision is made through TPM. The data is also stored in the IDSS database for future reference and decision support. This system ensures that teacher placement is more systematic, efficient, and based on data analysis.

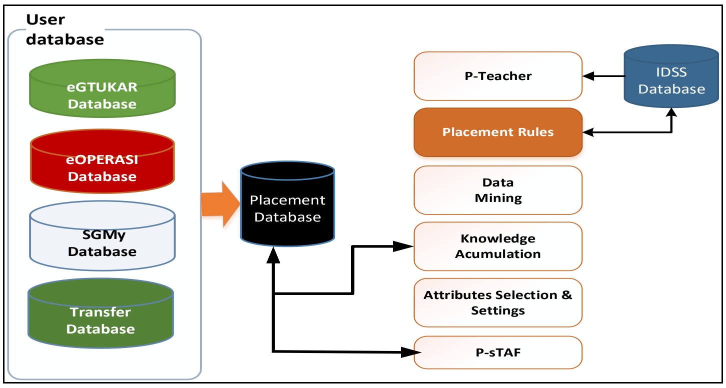

FIGURE 6: TPM Implementation in KPM

The placement attributes selection at KPM is to help placement specialists in matching the suitable placement attributes, which is to maximize the quality of decision making. Therefore, constraints that affect the placement process, such as established placement policies and regulations, personnel, job, human factors, and support factors must be taken into account in the selected combination of placement attributes. This involves the selection that includes the categories and components.

In particular, staff placement requirements were explored through expert interviews to validate and further identify key findings in advance from the literature review aiming to align key staffing needs with selected placement attributes. The study was framed by structured questions, after which, a customer satisfaction survey was based on the findings from the interviews.

Relevant variables of the domain were elicited by domain experts using structured interviews on reading teacher placements at KPM. To facilitate this process, the interview took place in person and online, the two experts reviewed the questions in advance of the interview, as a means to better understand the context of the questions. Ten specialists were interviewed who directly assist the researchers in teaching placement (Augustine et al., 2019).

4.2 Analysis of Placement Model

The Teacher Placement Prototype System Development is an important stage to determine the TPM effectiveness. To obtain an understanding of the TPM feasibility, the prototype was designed to process and analyze teacher placement data to ensure that predictions reflect what is needed in the workforce. Data-driven decision-making processes are integrated into the prototype.

The prototype is aimed to test the proposed predictive model for teacher placement and the role of data mining classification techniques in improving placement outcomes. The system is aimed to simulate real-world teacher placement scenarios and provide a structured approach to analysing and optimising placement decisions. Figure 7: Key Components in the Design Development of the Teacher Placement Decision Support System (TP-DSS).

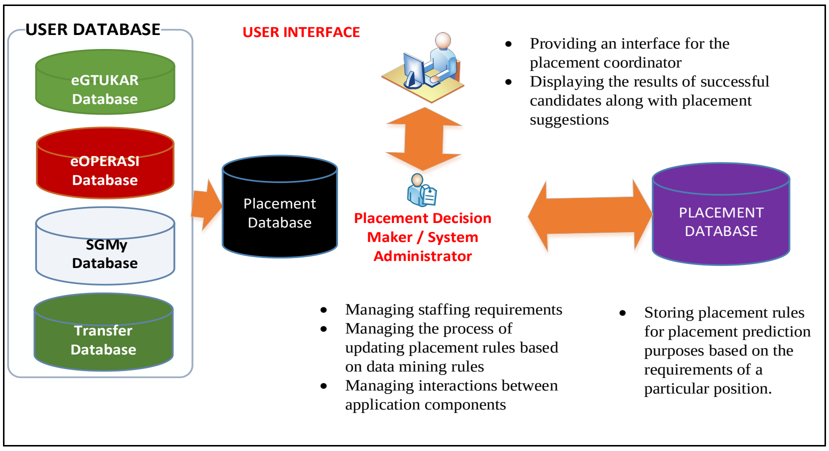

FIGURE 7 : Main components in the design development of the TP-DSS

The prototype was designed to enable prediction of teacher placement using historical data in an automated way. It employs classification algorithms to understand placement parameters including employment status, subject of specialization, location preferences, and work experience. The accuracy of the model was assessed by comparing predictions of placements with actual historical placements obtained from eGTUKAR system. This pretrained model enables the prototype to provide placement recommendations that are data-driven rather than relying on manual selection methods.

To verify unbiased placement recommendations, classification performance was validated with a separate test dataset during the prototype testing phase. The classification rules generated by the J48 algorithm represent instructions utilized during decision-making within the TP-DSS (Teachers Placement Decision Support System). The placement algorithm in the TP-DSS is driven by these rules to ensure that each decision is made based on systematic and structured data analysis. The effectiveness of the placement model was evaluated using 3,096 new application records from candidates in the second session of 2022. Placement model evaluation was conducted using a significance test (t-test) to determine if a particular classification technique is statistically significant compared to others. Previous studies suggest that decision tree classifiers often produce highly accurate models. In this study, analysis found that J48 decision tree classifier achieved slightly higher accuracy percentages than other classifiers. To confirm this observation for placement data, a paired t-test was performed.

The t-test is used to determine whether there is a significant difference between the means of two data groups. In this analysis, we conducted a t-test on the TPC and TPM columns in the given dataset. Table 15 The descriptive statistics for the TPC and TPM and Table 16 show t-Test Result.

Table 15 : Descriptive statistics for the TPC and TPM

Statistic	**TPC**	**TPM**
Sample size (n)	2,562	2,562
Mean	0.609	0.617
Standard deviation	0.488	0.486
Minimum	0.000	0.000
Maximum	1.000	1.000

Table 16 : t-Test Results

Test Statistic	**Value**
t-value	-0.631
p-value	0.528

Based on the t-test results, the obtained p-value is 0.528, which is greater than the significance level of 0.05. This means that there is no statistically significant difference between the means of TPC and TPM. Therefore, the null hypothesis stating that both means are equal cannot be rejected.

Based on this analysis, there is no statistical evidence to suggest a significant difference between the TPC and TPM values in this dataset. If necessary, further analysis using other statistical methods can be conducted to gain deeper insights.

DISCUSSION

The study successfully develops a TPM using data mining techniques to enhance the efficiency and fairness of teacher allocation. By analyzing historical placement data and applying classification techniques, the model identifies key factors influencing teacher assignments, reducing mismatches and improving workforce distribution.

The results indicate that the J48 technique achieved the highest accuracy (96.30%) compared to Decision Tree, Naïve Bayes, Random Forest, and K-Nearest Neighbors. The model's effectiveness was further evaluated using 3,096 new application records. The results showed that the proposed model approved 1,554 placements compared to 1,5421 in the actual decisions. The t-test results also showed a high p-value ($p > 0.05$), proving that the model's decisions were comparable to those made by the TPC with no significant difference.

Future research should enhance TPM by integrating workforce forecasting, financial planning, and training development. Expanding machine learning techniques, such as fuzzy logic and genetic algorithms, may further refine placement predictions. The model's automated implementation could link a knowledge transformation engine to a knowledge database for better efficiency. Improving the system prototype will also ensure its practical application in real-world teacher placement scenarios.

This study successfully achieves its objective of developing a structured TPM using data mining, contributing to more systematic and data-driven teacher allocation within the KPM. The model serves as a foundation for optimizing workforce distribution in education and can be adapted for broader human resource management applications.

In conclusion, the developed staff placement model can be effectively applied to develop a teacher placement model (TPM) at MOE. This TPM model has the potential to serve as an essential decision-support tool for MOE in improving the efficiency of teacher placement in the future.

REFERENCES

[1] Ahmad, R. & Siti, N. (2023). Pengurusan Data Penempatan Pensyarah: Cabaran dan Penyelesaian. Jurnal Pengurusan Pendidikan Malaysia.

[2] BPSH, K. (2024a). Aplikasi eGTUKAR (3.0; Issue Penempatan Guru, pp. 0–5). https://epgo.moe.gov.my

[3] BPSH, K. (2024b). eOperasi. KPM. https://eoperasi.moe.gov.my/

[4] Dai, J. (2023). Research on micro-decision mechanism of university teachers' mobility based on data mining. 12779(Icmir 2023), 18. https://doi.org/10.1117/12.2688703

[5] Encarnacion, R. E., & Cruz, M. E. L. T. (2022). Predicting Student's Internship Performance Using J48 Algorithm. 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), 1–6. https://doi.org/10.1109/ICECCME55909.2022.9987929

[6] Ensimau, N. K., Hamzah, M. I., Yasin, M. H. M., & Nasri, N. M. (2024). The Implementation of Zero Reject Policy in Malaysia: A systematic Review. Proceeding of International Conference on Special Education in South East Asia Region, 1(1), 243–260. https://doi.org/10.57142/picsar.v1i1.50

[7] Hilmi, M. A. (2022). The Impact of Sending Top College Graduates to Rural Primary Schools. Southeast Asian Economies, 39(S), S62–S79. https://doi.org/10.1355/ae39-se

[8] HRMIS. (2024). HRMIS - Laman Rasmi HRMIS. https://www.eghrmis.gov.my

[9] Hu, Y. (2022). College Teacher Training Resource Management System Based on Data Mining Technology. 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), 1321–1324. https://doi.org/10.1109/IPEC54454.2022.9777383

[10] Hudson, B., Hunter, D., & Peckham, S. (2019). Policy failure and the policy-implementation gap: can policy support programs help? Policy Design and Practice, 2(1), 1–14. https://doi.org/10.1080/25741292.2018.1540378

[11] Langeveldt, D. C., & Pietersen, D. (2024). Data-Driven Strategies for Addressing Challenges in Teacher Placement: A Legal and Pedagogical Analysis for Inclusive Education in South Africa. E-Journal of Humanities, Arts and Social Sciences, 4(12), 219–229. https://doi.org/10.38159/ehass.202341219

[12] Lei, J. (2020). Cross-Validation With Confidence. Journal of the American Statistical Association, 115(532), 1978–1997. https://doi.org/10.1080/01621459.2019.1672556

[13] Lim, K. & Hassan, M. (2024). Pendekatan Teknologi dalam Penempatan Pensyarah: Kajian Kes di Universiti Tempatan. Jurnal Teknologi Pendidikan.

[14] Liu, Y. (2020). Analysis of Human Resource Management Mode and Its Selection Factors Based on Decision Tree Algorithm. 2020 International Conference on Advance in Ambient Computing and Intelligence (ICAACI), 139–142. https://doi.org/10.1109/ICAACI50733.2020.00035

[15] Manap, H. R. (2024). The Influence of Selection and Assignment Placement on the Quality of Human Resources in the Education Sector.". INTERDISIPLIN Journal of Qualitative and Quantitative Research, 1(2), 80–85. https://doi.org/https://doi.org/10.61166/interdisiplin.v1i2.16.

[16] Mary-Magdalene, W., Flora, C., & Scholastica, W. A. (2023). Teacher placement and its impact on students performance: The perception of teachers at Junior High Schools of the Kassena/Nankana Municipality. Educational Research and Reviews, 18(11), 363–372. https://doi.org/10.5897/err2023.4342

[17] Nayeri, Mahmoud Dehghan, Moein Khazaei, and D. A. (2022). The Drivers of Success in New-Service Development: Rough Set Theory Approach. International Journal of Services and Operations Management, 43 (4), 421. https://doi.org/https://doi.org/10.1504/ijsom.2022.127465

[18] Ojokoh, B., Akinsulire, V., & Isinkaye, F. (2019). An Automated Implementation of Academic Staff Performance Evaluation System based on Rough Sets Theory. Australasian Journal of Information Systems, 23, 1–20. https://doi.org/10.3127/ajis.v23i0.2033

[19] Shamsul. (2023). Model Penempatan Staf Bagi Kerangka Kerja Sistem Sokongan Pintar Berasaskan Perlombongan Data. Universiti Kebangsaan Malaysia.

[20] Sriwindono, H., Rosa, P. P., & Pinaryanto, K. (2022). Teacher Placement using K-Means Clustering and Genetic Algorithm. Conference Series, 4(January), 43–51. https://doi.org/10.34306/conferenceseries.v4i1.669

[21] Tasfia, S., Reno, S., Jahan, N., & Al Mamun, A. (2024). Exploring Weka and Python for Educational Data Mining: Naïve Bayes vs. J48. 2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), 470–475. https://doi.org/10.1109/ICMCSI61536.2024.00074

[22] Wei, F. (2022). Performance Evaluation of Tourism Human Resource Management Based on Fuzzy Data Mining. Journal of Mathematics, 2022(1), 3745377. https://doi.org/https://doi.org/10.1155/2022/3745377

[23] Wu, Y., Wang, Z., & Wang, S. (2021). Human Resource Allocation Based on Fuzzy Data Mining Algorithm. Complexity, 2021(1), 9489114. https://doi.org/https://doi.org/10.1155/2021/9489114

[24] Xu, K. (2023). Application of improved association rules algorithm and cloud service system in human resource management. International Journal of System Assurance Engineering and Management. https://doi.org/10.1007/s13198-023-02175-w

[25] Yan, X., Deng, X., & Sun, S. (2020). Analysis and Simulation of the Early Warning Model for Human Resource Management Risk Based on the BP Neural Network. Complexity, 2020(1), 8838468. https://doi.org/https://doi.org/10.1155/2020/8838468

[26] Zhang, C. (2022). Evaluation and analysis of human resource management mode and its talent screening factors based on decision tree algorithm. The Journal of Supercomputing, 78(13), 15681–15713. https://doi.org/10.1007/s11227-022-04499-z

[27] Zhao, B., Xu, Y., & Cheng, J. (2021). Evaluation and Image Analysis of Enterprise Human Resource Management Based on the Simulated Annealing-Optimized BP Neural Network. Computational Intelligence and Neuroscience, 2021(1), 3133065. https://doi.org/https://doi.org/10.1155/2021/3133065

[28] Zhou, Y. (2023). Development and Training Strategies of College Teachers Based on Data Mining Technology. Mobile Information Systems, 2023(1), 7103391. https://doi.org/https://doi.org/10.1155/2023/7103391