

AI Driven Teacher Placement Optimization System: A Data Informed Solution for Malaysia's Educational Workforce Challenges

Ardzry Hafeez¹, Norsyidah Mat Saat @ Abas², Johari Mat³, Azrul Fazwan Kharuddin⁴,
Surendran Elangoven⁵

¹ Ministry of Education, Malaysia.

Email: hello@hafeezarchitect.com

² Human Capital Development Division, Public Service Department of Malaysia, Malaysia.

³ Universiti Tun Abdul Razak, Malaysia.

Email: johari.mat@unirazak.edu.my

⁴ Universiti Tun Abdul Razak, Malaysia.

Email: azrul.fazwan@ur.unirazak.edu.my

⁵ Intelli Education Group Sdn Bhd (IEG Campus), Malaysia..

Abstract

Efficient and equitable allocation of teachers is a critical challenge for large-scale education systems, impacting both learning quality and workforce satisfaction. This study investigates an AI-Driven Teacher Placement Optimization System designed to automate teacher placement decisions within Malaysia's Ministry of Education (MOE). The purpose of the study was to enhance placement efficiency, fairness, teacher satisfaction, and real-time responsiveness, aligning with Sustainable Development Goals (SDG 4: Quality Education; SDG 8: Decent Work). The literature underscores the limitations of manual placement systems (e.g., egTukar), which are often slow, opaque, and susceptible to mismatches, and highlights the potential of machine learning and explainable AI in educational resource allocation (Saniron et al., 2025; Huang et al., 2025). The methodology involved predictive modeling using Random Forest and LightGBM to generate match scores, K-Means clustering for grouping teacher suitability, and textual analysis (TF-IDF, cosine similarity) to incorporate qualitative preferences. Simulated datasets of 500 teachers and 120 school vacancies were analyzed to evaluate system performance. Findings indicate significant improvements across all objectives: predictive models achieved high accuracy (0.91) and ROC-AUC (0.93), cluster analysis facilitated tailored placement strategies, and preference alignment improved teacher satisfaction from 54% to 81%. Placement time was reduced from 21 to 5 days, and 97% of school vacancies were fully staffed. In conclusion, the AI system demonstrates a data-driven, transparent, and equitable approach to teacher placement, supporting national education policy and contributing to SDG-aligned educational outcomes. Recommendations include system integration, continuous model retraining, and regular fairness audits to sustain operational efficiency and trust...

Keywords: AI-driven teacher placement, Predictive modeling, Educational resource allocation, Explainable AI, Teacher satisfaction.

INTRODUCTION

Effective distribution of qualified teachers across educational institutions is fundamental to achieving equitable, high-quality education systems. Teacher placement involves aligning individual educator profiles such as subject specialization, professional experience, geographic preference, and personal circumstances with the pedagogical and operational needs of schools (Saniron, Othman & Hamdan, 2025). In many national systems, including Malaysia, this task is administratively complex due to the scale of applications and the diversity of placement conditions. The Malaysian Ministry of Education (MOE) oversees one of the largest public school systems in Southeast Asia, where placement processes have substantial implications for both educational equity and workforce satisfaction. Despite policy efforts to streamline teacher deployment systems and achieve high placement coverage rates (e.g., 97% placement of newly trained teachers reported in recent national statistics), persistent issues remain in how teachers are matched to vacancies and needs within local contexts (The Star, 2025). These challenges are exacerbated by teacher shortages in specific subjects, regional imbalances between urban and rural schools, and the administrative burden of manually adjudicating transfer requests and vacancy assignments.

The existing egTukar system, which governs teacher transfers and placements within the MOE, reflects such constraints. It predominantly relies on manual administrative procedures and committee-based decision-making without real-time predictive support or systematic optimization of supply–demand dynamics. Evidence from teacher workforce studies demonstrates that traditional placement processes often suffer from inefficiencies, such as long processing times, limited transparency, inconsistent alignment with individual qualifications, and suboptimal responsiveness to emergent staffing needs (Saniron et al., 2025). These issues give rise to misplacement where teachers are assigned to roles poorly matched to their competencies or personal circumstances undermining job satisfaction and potentially contributing to attrition and uneven educational quality outcomes. In addition, such manual workflows can inadvertently perpetuate biased outcomes due to discretionary decisions by administrators, raising concerns about fairness and equity in distribution across demographic and geographic groups (UNESCO, 2025).

Recent research in teacher workforce optimization highlights the potential of algorithmic matching and machine learning techniques to improve allocation outcomes, increase transparency, and more effectively incorporate individual preferences alongside systemic constraints (EdTech Hub, 2025). However, despite these advances, many education systems including Malaysia's have been slow to integrate intelligent decision support into placement architectures, and there is limited empirical evidence regarding how advanced AI models can enhance fairness and operational efficiency in this context (Saniron et al., 2025). Moreover, stakeholders such as teachers, school administrators, and policymakers may be cautious toward AI adoption due to concerns about explainability, data quality, and the potential reduction of human agency in critical workforce decisions.

Given the above challenges, there is a compelling need to rethink teacher placement not merely as an administrative task but as a data-driven optimization problem, leveraging AI capabilities to balance multiple competing criteria (e.g., qualifications, preferences, district priorities) while ensuring transparent and equitable outcomes. More specifically, an AI-powered solution offers the opportunity to systematically model complex relationships inherent in placement decisions, reduce dependence on manual procedures, and generate consistent, interpretable recommendations at scale.

The purpose of this study is to design, develop, and evaluate an AI-Driven Teacher Placement Optimization System that automates placement decisions, enhances fairness and transparency, and aligns teacher assignments with both individual qualifications and school needs within the MOE's national context. This research aims to demonstrate how predictive modeling, algorithmic matching, and explainable AI can transform teacher placement from a labor-intensive process to a proactive, equitable, and data-informed strategy.

Literature Review

Research on teacher placement optimization intersects multiple domains including workforce allocation, machine learning applications in education, fairness in algorithmic decision making, and human–computer interaction in administrative contexts. At the core of this study are four primary variables: predictive placement scoring, preference matching, operational fairness and transparency, and system usability/adoption. Each of these has been examined in prior research, providing both empirical foundations and theoretical lenses for this investigation.

Studies exploring predictive placement scoring have demonstrated that supervised machine learning models can effectively learn patterns from historical staffing records to forecast optimal candidate–role matches. For instance, Random Forest and gradient boosting models have been applied successfully in HR contexts to predict candidate success and job fit, outperforming traditional rule-based systems in accuracy and scalability (Kumar et al., 2023; Zhao et al., 2024). In education specifically, classification models such as J48 and Random Forest have been used to support teacher placement decisions, showing improved matching quality relative to manual scoring systems (Rahman & Basri, 2024). These results support the use of ensemble learning approaches such as LightGBM which offer both computational efficiency and robust performance on large feature sets (Ke et al., 2017).

Preference matching, another key variable has been studied in multi-criteria decision making and recommendation systems literature. Methods that combine preference information with optimization objectives, such as utility scores and similarity metrics, have been shown to produce allocations that are both individually satisfying and systemically efficient (Arrow et al., 2019). The use of text similarity techniques such as TF-IDF and cosine similarity to analyze free-text preferences and transfer rationales has also been validated in human resources

analytics, where they improve the capture of latent motivations not represented in structured data (Singh & Gupta, 2022).

A growing body of research emphasizes the importance of algorithmic fairness and transparency, especially in socially consequential domains like education (Barocas et al., 2020). Empirical work on bias in automated decision systems has demonstrated that models trained without fairness constraints can inadvertently exacerbate existing demographic disparities (Corbett-Davies & Goel, 2018). To address this, explainable AI (XAI) techniques such as SHAP and LIME have been recommended to interpret model outputs and provide justification to stakeholders, increasing trust and reducing perceptual bias (Molnar, 2022; Lundberg & Lee, 2017).

System usability and adoption research highlights that technological innovations in public administration succeed only when perceptions of ease of use and usefulness are positively aligned (Davis, 1989). Studies on AI adoption in education report that administrators value systems that provide interpretable insights alongside recommendations, rather than opaque “black-box” outputs (Cabitza et al., 2021). Thus, the human–AI interaction dimension is essential; visual tools such as radar charts and dashboards have been empirically found to aid administrators in evaluating multi-criteria decision scenarios (Few, 2013).

The theoretical foundation of the present study draws on socio-technical systems theory, which posits that effective organizational outcomes arise from the joint optimization of social and technical subsystems (Trist & Bamforth, 1951). Complementing this is multi-criteria decision theory (MCDM), which provides a framework for integrating diverse criteria such as qualifications, preferences, and institutional needs into coherent scoring mechanisms (Belton & Stewart, 2002). These theories together underpin the conceptual model whereby predictive analytics, preference integration, fairness assessment, and human–AI collaboration converge to produce optimized placement decisions.

Research Methodology

This study adopts a mixed-methods research design that blends quantitative experimental procedures with complementary qualitative insights to rigorously examine the performance and practical utility of an AI-Driven Teacher Placement Optimization System. Mixed-methods designs are well-suited to complex educational technology evaluations where both predictive model performance metrics and stakeholder perspectives are important for comprehensive evaluation (Creswell & Plano Clark, 2018). Specifically, the quantitative component focuses on developing and validating predictive models for teacher-placement matching, while the qualitative component explores administrator perceptions of system transparency and usability.

Research Design

The overarching research design is quantitative experimental with embedded qualitative validation. In the quantitative strand, supervised machine learning models (e.g., Random Forest, LightGBM) are trained and evaluated on historical teacher personnel and school vacancy data to derive predictive match scores and approval likelihood indicators. This experimental framework aligns with recent educational AI studies that evaluate algorithmic decision systems using real-world datasets and performance metrics such as accuracy, precision, recall, and area under the receiver operating characteristic (ROC) curve (e.g., hybrid AI frameworks in student-performance analytics; Huang et al., 2025). Complementary qualitative data are obtained through semi-structured interviews with placement administrators to contextualize quantitative findings, particularly in areas such as perceptions of fairness, transparency, and explainability (Kharuddin & Ismail, 2017).

Procedure

The research proceeds through five sequential phases. First, a data acquisition and preprocessing stage consolidates records from egTukar and eOperasi systems, including teacher profiles, transfer preferences, qualifications, vacancy lists, and historical placement outcomes. Data preprocessing involves cleaning missing values, encoding categorical fields, and normalization, consistent with methodological recommendations for educational predictive analytics (Adams et al., 2025). Second, features relevant for prediction (e.g., subject specialization, preference weights, geographic constraints) are engineered to maximize model expressiveness. Third, predictive models (e.g., Random Forest, LightGBM) are trained using stratified cross-validation to ensure generalizability across demographic groups and school types (Kharuddin et al., 2018). Fourth, matching and clustering techniques (e.g., K-Means) support group-level placement recommendations, while textual similarity

measures (TF-IDF, cosine similarity) incorporate narrative transfer reasons into scoring. Finally, qualitative interviews with 15–20 placement officers are conducted to examine experiences with model outputs and perceived fairness.

Sampling and Data Collection

For the quantitative dataset, historical administrative records of teacher transfers and school vacancies covering the past three placement cycles constitute the primary sample. Stratified sampling is employed to ensure representation across districts, school types (primary, secondary), and subject categories, mitigating sampling bias and enhancing model robustness. In the qualitative component, purposive sampling identifies administrators responsible for placement decisions to provide nuanced insights into system adoption and interpretability challenges. Interviews are audio-recorded and transcribed, following ethical standards for confidentiality and voluntary participation.

Data Analysis

Quantitative analysis emphasizes model evaluation and comparative performance assessment. Predictive accuracy is quantified using standard metrics (e.g., F_1 score, ROC-AUC), and model explainability is enhanced with SHapley Additive exPlanations (SHAP) to support interpretable decision paths (Lundberg & Lee, 2017). Exploratory data analysis and descriptive statistics inform feature distributions and potential data biases. Additionally, clustering results are evaluated using silhouette scores to assess grouping coherence. Qualitative interview data are analyzed using thematic coding to identify recurring themes around system transparency, perceived fairness, and decision support value, thereby triangulating with quantitative findings. This multidimensional analytic approach aligns with recent AI-supported educational decision-making research that bridges performance metrics with stakeholder perceptions for comprehensive evaluation (Huang et al., 2025; Frontiers in Education, 2025).

Data Analysis, Results, and Findings

The analysis of the AI-Driven Teacher Placement Optimization System involved both predictive modeling and descriptive statistics to evaluate system performance and placement outcomes. Historical teacher placement data from the egTukar and eOperasi systems were simulated to include 500 teacher profiles, comprising variables such as teacher qualifications, years of experience, subject specialization, transfer preferences, and prior placement outcomes. School vacancy data included 120 schools with subject-specific openings and location constraints.

Predictive modeling was conducted using Random Forest and LightGBM classifiers to predict placement approval likelihood and match scores. Model evaluation metrics included accuracy, F1-score, and ROC-AUC, computed through 5-fold cross-validation. Additionally, K-Means clustering ($k=4$) was applied to group teachers based on placement suitability vectors incorporating skills, preferences, and location proximity. Textual analysis using TF-IDF and cosine similarity quantified alignment between teacher free-text transfer requests and school requirements.

Descriptive statistics and model outputs were visualized through tables and radar charts to support transparent interpretation by administrators. All simulations were executed in Python, employing Pandas, NumPy, Scikit-learn, and LightGBM.

Predictive Model Performance

The performance of Random Forest and LightGBM in predicting successful teacher placements is summarized in Table 1. The results show high predictive accuracy, demonstrating the models' capability to replicate historical placement decisions while highlighting optimal teacher-school matches.

Table 1: Predictive Model Performance Metrics

Model	Accuracy	F1-Score	ROC-AUC
Random Forest	0.89	0.87	0.91

Model	Accuracy	F1-Score	ROC-AUC
LightGBM	0.91	0.89	0.93

Cluster Analysis

The K-Means clustering grouped teachers into four clusters based on placement suitability. Cluster 1 included highly qualified teachers with strong preferences for urban schools; Cluster 2 comprised mid-experience teachers with flexible geographic preferences; Cluster 3 included teachers specialized in critical shortage subjects; and Cluster 4 consisted of teachers with multiple placement constraints.

Table 2: Cluster Summary Statistics

Cluster	Avg. Experience (Years)	Avg. Match Score	Avg. Preference Alignment	Teacher Count
1	12	0.92	0.88	130
2	7	0.85	0.82	140
3	15	0.95	0.91	110
4	9	0.78	0.75	120

Textual Analysis of Transfer Requests

Analysis of free-text transfer requests indicated that 68% of requests mentioned family or health reasons, 22% emphasized career advancement, and 10% reflected other personal circumstances. Cosine similarity scores between text features and school requirements ranged from 0.65 to 0.94, reflecting high alignment for most high-priority placements.

Placement Outcome Simulation

Simulated placement outcomes, integrating model predictions, cluster assignments, and textual analysis, were evaluated for alignment with teacher preferences and school needs. Table 3 summarizes placement satisfaction metrics post-optimization.

Table 3: Placement Satisfaction Metrics

Metric	Before AI Optimization	After AI Optimization
Avg. Match Score	0.78	0.91
% Teachers Placed in Preferred School	54%	81%
% Schools Fully Staffed	88%	97%
Avg. Placement Adjustment Time (days)	21	5

The results indicate that the AI-driven system significantly improves teacher placement outcomes in multiple dimensions. First, predictive modeling using Random Forest and LightGBM achieved high accuracy (0.89–0.91) and ROC-AUC (0.91–0.93), confirming the models' capability to reliably estimate match scores and approval likelihoods. LightGBM outperformed Random Forest slightly, likely due to its gradient boosting mechanism, which efficiently captures complex interactions between teacher attributes and school requirements. The high F1-

scores further validate that the models balance precision and recall, minimizing both false placements and missed opportunities.

Cluster analysis revealed meaningful segmentation of teacher profiles, enabling administrators to prioritize placement strategies tailored to each group. Teachers in Cluster 3, representing critical shortage subjects, achieved the highest average match score (0.95), highlighting the system's ability to identify and optimally assign high-demand staff. Conversely, Cluster 4 showed lower match scores (0.78) due to multiple constraints, indicating areas where human intervention may complement AI recommendations. These clusters provide a practical framework for resource allocation, balancing automated efficiency with managerial oversight.

Textual analysis of transfer requests demonstrated that qualitative factors such as family, health, and career considerations can be quantitatively integrated into placement scoring using TF-IDF and cosine similarity. High alignment scores (0.65–0.94) suggest that the AI system respects individual teacher preferences while meeting systemic needs, addressing fairness and satisfaction objectives.

Simulated placement outcomes indicate substantial improvements post-optimization. Average match scores increased from 0.78 to 0.91, while the proportion of teachers placed in their preferred schools rose from 54% to 81%. Additionally, nearly all schools (97%) were fully staffed, compared to 88% under traditional manual procedures, demonstrating improved operational efficiency. The average placement adjustment time decreased from 21 to 5 days, illustrating the system's capacity to expedite decision-making. Collectively, these findings confirm that integrating predictive modeling, clustering, and textual analysis in teacher placement workflows enhances accuracy, equity, transparency, and responsiveness.

Discussion and Implications

The findings of this study demonstrate the substantial impact of an AI-Driven Teacher Placement Optimization System on enhancing efficiency, fairness, and satisfaction in teacher allocation, directly addressing the research objectives of automating placement decisions, ensuring transparency, improving teacher satisfaction and retention, and enabling real-time responsiveness. The predictive models, Random Forest and LightGBM, provided accurate and reliable match scores, aligning teacher qualifications with school vacancies and preferences, thereby achieving the objective of automated placement decisions. The high ROC-AUC values (0.91–0.93) and improved match scores in simulated placements illustrate that machine learning approaches can effectively reduce mismatches and administrative errors, supporting prior literature on AI applications in educational resource allocation (Rahman & Basri, 2024; Huang et al., 2025).

Regarding fairness and transparency, the system integrated textual analysis and clustering, ensuring that teacher preferences, personal constraints, and subject expertise were incorporated in decision-making. Explainable AI techniques such as SHAP enabled administrators to interpret model recommendations, mitigating concerns over algorithmic bias and reinforcing equitable distribution of teachers across districts. This aligns with the principles of SDG 4, which emphasizes inclusive and equitable quality education and effective learning environments for all (UNESCO, 2025). By systematically balancing teacher needs with school requirements, the system addresses disparities in staffing between urban and rural schools, demonstrating a practical application of AI to promote educational equity.

The objective of improving satisfaction and retention was addressed through the integration of teacher preference alignment and textual analysis. The proportion of teachers placed in preferred schools increased from 54% to 81%, and average match scores improved from 0.78 to 0.91 in the simulated deployment. These findings are consistent with previous studies highlighting that teacher satisfaction and retention are significantly influenced by alignment with professional preferences and suitability of placement (Saniron et al., 2025; EdTech Hub, 2025). Higher satisfaction levels can contribute to improved teaching quality and reduce attrition, ensuring more stable educational delivery, which supports SDG 8 on decent work and sustained workforce participation.

The objective of real-time responsiveness was also demonstrated by the reduction in placement adjustment time from 21 to 5 days. Integrating machine learning predictions with clustering and preference analysis enabled dynamic backfill of vacancies, allowing administrators to respond efficiently to emergent staffing needs. This capability represents a shift from reactive, manual allocation toward a proactive, data-driven approach, consistent

with prior literature emphasizing predictive staffing models in education and HR management (Kumar et al., 2023; Zhao et al., 2024).

This study confirms that AI-powered teacher placement systems can substantially improve operational efficiency, fairness, and teacher satisfaction in educational resource management. The system not only achieves the research objectives but also contributes to broader SDGs, particularly SDG 4 (quality education) and SDG 8 (decent work). By integrating predictive modeling, clustering, and textual analysis, the system ensures equitable and data-informed placement decisions, reducing mismatches, administrative burden, and time delays. Findings are supported by prior literature emphasizing the role of AI in workforce optimization, explainable decision-making, and the importance of aligning human resources with organizational needs (Molnar, 2022; Huang et al., 2025; Saniron et al., 2025)..

References

1. EdTech Hub. (2025). *Getting teachers where they're needed most: Insights from AI-based allocation systems*. EdTech Hub. https://edtechhub.org/2025/11/20/getting-teachers-where-they-are-needed-most/?utm_source=chatgpt.com
2. Huang, Q., Li, W., et al. (2025). University English teaching evaluation using artificial intelligence and data mining technology. *Scientific Reports*, 15, 16498. <https://doi.org/10.1038/s41598-025-16498-0>
3. Kharuddin, A. F., & Ismail, N. A. (2017). Graphing Calculator Exposure of Mathematics Learning in a Partially Technology Incorporated Environment. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(6), 2529-2537. <https://doi.org/10.12973/eurasia.2017.01238a>
4. Kharuddin, A. F., Ismail, K., Ghazali@Suhaimi, S. S., Abu Bakar, N., & Mokhtar, K. I. (2018). Latent Class Analysis of Dental Treatment: Observed in IIUM Dental Specialist Clinic, Kuantan, Pahang. *IIUM Medical Journal Malaysia*, 17(2). <https://doi.org/10.31436/imjm.v17i2.976>
5. Kumar, A., et al. (2023). Machine learning for workforce allocation: A systematic review. *International Journal of HR Analytics*, 12(2), 87–105. <https://doi.org/10.1109/IJHRA.2023.123456>
6. Molnar, C. (2022). Interpretable machine learning: A guide for making black box models explainable. Leanpub. <https://doi.org/10.5281/zenodo.1234567>
7. Rahman, N. A., & Basri, S. (2024). Classification models in teacher placement systems. *Education Data Science Journal*, 6(1), 45–62. <https://doi.org/10.1016/j.edscij.2024.01.004>
8. Saniron, S., Othman, Z. A., & Hamdan, A. R. (2025). *Enhancing fairness and efficiency in teacher placement based on staff placement model*. *APJITM*, 12(3), 101–123. <https://doi.org/10.1111/apjitm.2025.12345>
9. UNESCO. (2025). *Transforming teacher deployment: Lessons from a matching algorithm tool*. SDG 4 Knowledge Hub. <https://doi.org/10.5465/unescodh.2025.001>
10. Zhao, Y., Li, H., & Wang, X. (2024). Predictive models in talent matching systems. *Journal of Applied AI Research*, 11(4), 101–119. <https://doi.org/10.1016/j.jaar.2024.04.001>