A Machine Learning Framework for Human Resource Management in Higher Education: A Case Study of Nigerian Universities

Onwuegbuchulam Gideon[1]; Ugah John[2];
Chukwuemeka Anyim[3]; Wilfred Adigwe[4]

[1] Department of Computer Science, Ebonyi State University, Nigeria
[2] Department of Computer Science, Ebonyi State University, Nigeria
[3] David Umahi Federal University of Health Sciences, Nigeria
[4] Department of Computer Science, Southern Delta University, Nigeria

Abstract
Effective management of academic staff is central to the productivity of higher education institutions, yet traditional Human Resource Management (HRM) in Nigerian universities remains manual, subjective, and inefficient. This study proposes a Machine Learning (ML)-based HRM framework to enhance performance evaluation, workload balancing, and promotion recommendations. Using the CRISP-DM methodology, datasets on teaching load, research output, administrative roles, and professional development were collected and preprocessed. Three supervised ML algorithms—Random Forest, Support Vector Machines (SVM), and XGBoost—were trained and evaluated. Results show that the framework improves accuracy and consistency in staff assessments while reducing biases, with XGBoost outperforming other models in workload optimization and promotion prediction. The framework also features a scalable web-based architecture built with Flask and SQLAlchemy, supporting real-time analytics for administrators. This study presents a novel ML-driven HRM framework that enhances efficiency, fairness, and decision-making in resource-constrained academic environments.

Keywords: Human Resource Management (HRM); Machine Learning (ML); Nigerian universities; Academic staff evaluation; Workload balancing; Promotion recommendations

I Introduction

Human Resource Management (HRM) is central to university effectiveness since the quality of academic staff shapes teaching, research, and institutional reputation. Modern HRM increasingly relies on analytics to guide performance management, promotion, workload allocation, and staff development (Armstrong & Taylor, 2020; Davenport, Harris, & Shapiro, 2010). Globally, "people analytics" has improved fairness and efficiency in staffing decisions (Bock, 2015; Tursunbayeva, Di Lauro, & Pagliari, 2018). For universities in developing contexts such as Nigeria—where governance and resource challenges persist—data-informed HR processes are vital to sustain competitiveness (Saint, Hartnett, & Strassner, 2003).

Yet many Nigerian universities still depend on manual records and subjective judgments in appraisals, promotions, and workload allocation. These practices often enable favoritism, delay career progression, and create workload imbalances, which undermine morale and productivity. The lack of predictive analytics further prevents proactive planning (Saint et al., 2003; Davenport et al., 2010; Tursunbayeva et al., 2018). Concerns about fairness are heightened by the absence of bias auditing and explainability in decision-making (Raghavan, Barocas, Kleinberg, & Levy, 2020).

This study introduces a machine learning (ML)driven HRM framework tailored to Nigerian higher education. The framework

DOI: https://doi.org/10.5281/zenodo.17113277

applies supervised ML to automate performance evaluation, optimize workloads, and recommend promotion decisions in a more transparent and objective manner (Davenport et al., 2010; Tursunbayeva et al., 2018). Specifically, the study contributes by comparing Random Forests, Support Vector Machines (SVM), and XGBoost for predicting lecturer performance (Breiman, 2001; Cortes & Vapnik, 1995; Chen & Guestrin, 2016). It also demonstrates the application of the CRISP-DM methodology to ensure methodological rigor (Wirth & Hipp, 2000) and presents a modular web-based system with dashboards for administrators and lecturers. By addressing challenges such as delayed promotions and uneven workload distribution while incorporating fairness and explainability, the study shows how ML-enabled HRM can modernize academic administration in resource-constrained environments (Saint et al., 2003; Raghavan et al., 2020).

II Literature Review

AI and Machine Learning (ML) are transforming Human Resource Management (HRM) through automation, predictive analytics, and evidence-based decision-making. Unlike traditional HR practices marked by subjectivity and inefficiency, AI systems now support recruitment (Nawaz & Gomes, 2020), employee appraisal (Sharma & Sharma, 2020), and talent retention (Tursunbayeva et al., 2018). Models such as Random Forest, Support Vector Machines (SVM), and Gradient Boosting have been effective in classifying performance and forecasting progression (Zhang et al., 2020), making them valuable for higher education, where large volumes of staff data require objective analysis.

Globally, AI-driven HR systems show promising outcomes. IBM's Watson Talent Framework enhances recruitment and workforce planning through predictive insights (Guenole, Ferrar, & Feinzig, 2017), while Google's People Analytics demonstrates how ML can improve retention and performance (Bock, 2015). In academia, Nigerian studies have explored predictive appraisal systems: Oyelade, Ezugwu, and Oladipo (2020) achieved 85% accuracy with Naïve Bayes but lacked scalability; Afolabi, Olanrewaju, and Salami (2019) proposed a rule-based model with limited adaptability; and Mahmud, Alam, and Hasan (2018) developed a decision support tool for promotions that lacked predictive intelligence. These efforts highlight progress but also underscore challenges in adaptability and comprehensive integration.

Despite these advances, technical gaps remain. Many systems lack scalability across institutions (Oyelade et al., 2020), rely on static or rule-based approaches without predictive learning (Afolabi et al., 2019), or fail to incorporate fairness and bias auditing (Raghavan, Barocas, Kleinberg, & Levy, 2020). Thus, there is a need for a scalable ML-driven HRM framework that integrates predictive modeling, fairness considerations, and web accessibility. This study addresses these gaps by applying supervised learning models within the CRISP-DM methodology and designing a scalable system suited to Nigerian universities.

III Methodology

This study applied the Cross-Industry Standard Process for Data Mining (CRISP-DM) as the guiding framework. CRISP-DM provides a structured, iterative approach across six stages—business understanding, data understanding, preparation, modeling, evaluation, and deployment (Wirth & Hipp, 2000). Within this research, it was used to translate HR challenges in Nigerian universities—such as biased evaluation, workload imbalance, and promotion delays—into data-driven solutions, with iterations enabling refinement of models and system design for contextual relevance.Figure 1shows the CRISP-DM process adapted to HRM.

Data was drawn from institutional staff records, including teaching load, research output, administrative roles, professional development, and years of service. Preprocessing involved data cleaning to remove incomplete or duplicate records,

DOI: https://doi.org/10.5281/zenodo.17113277

normalization of variables, feature engineering (e.g., teaching-to-research ratios), and encoding of categorical variables into numerical form. The dataset was divided into training (70%) and testing (30%) subsets for model evaluation.Table 1ummarizes collected variables, types, and preprocessing steps.

Three supervised algorithms were used. Random Forest (RF) aggregates decision trees to handle high-dimensional data and reduce overfitting (Breiman, 2001). Support Vector Machines (SVM) classify data using optimal hyperplanes, with kernel functions capturing non-linear relationships (Cortes & Vapnik, 1995). Extreme Gradient Boosting (XGBoost) offers high accuracy and efficiency in structured data tasks, making it effective for promotion and workload predictions (Chen & Guestrin, 2016). These algorithms were selected for their robustness, interpretability, and suitability for decision-support.

The HRM system was developed as a modular, web-based application. Flask was used for the backend due to its lightweight integration with ML models, while SQLAlchemy managed database interactions. Object-Oriented Programming (OOP) principles enabled encapsulation of HR functions into modules for evaluation, workload distribution, and promotion. An interactive dashboard provided real-time visualization of metrics and predictive outcomes for administrators and lecturers. This architecture ensures scalability, usability, and adaptability across diverse university contexts.Figure 3 shows the system architectural diagram.

Figure 1: CRISP-DM process adapted to HRM

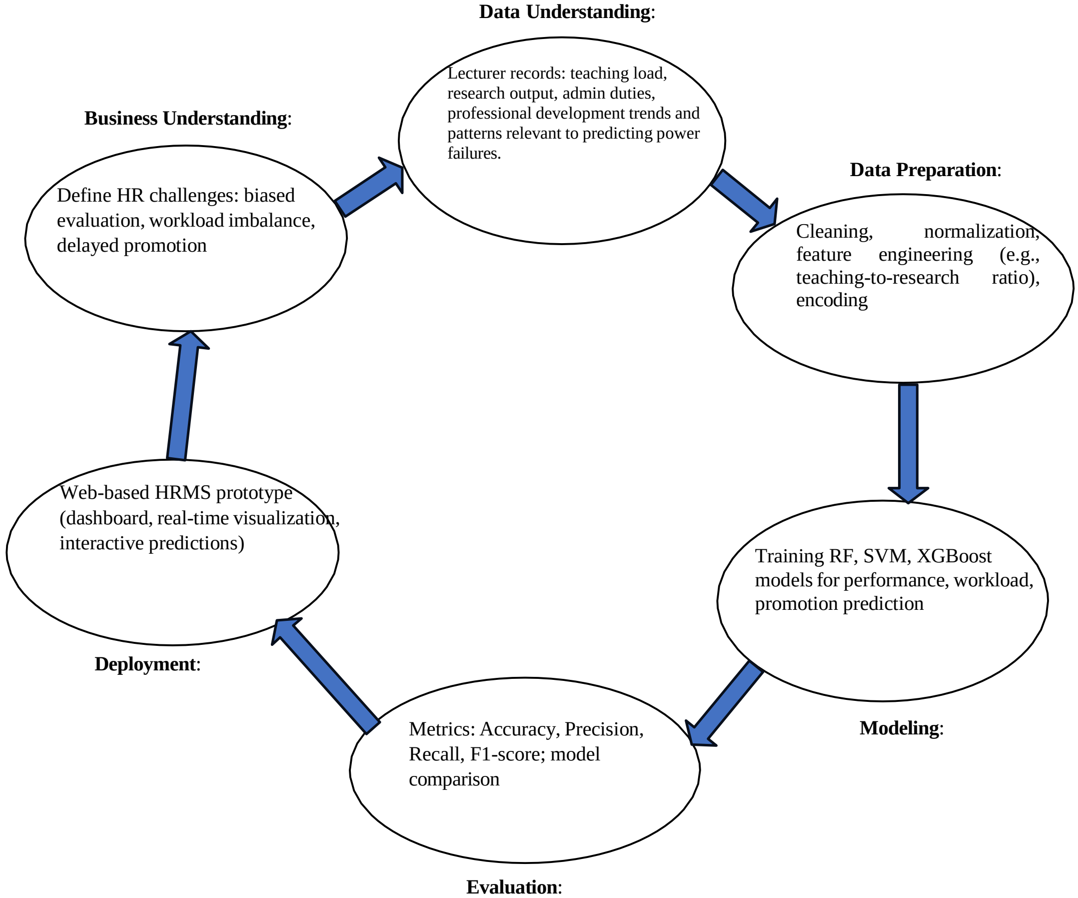
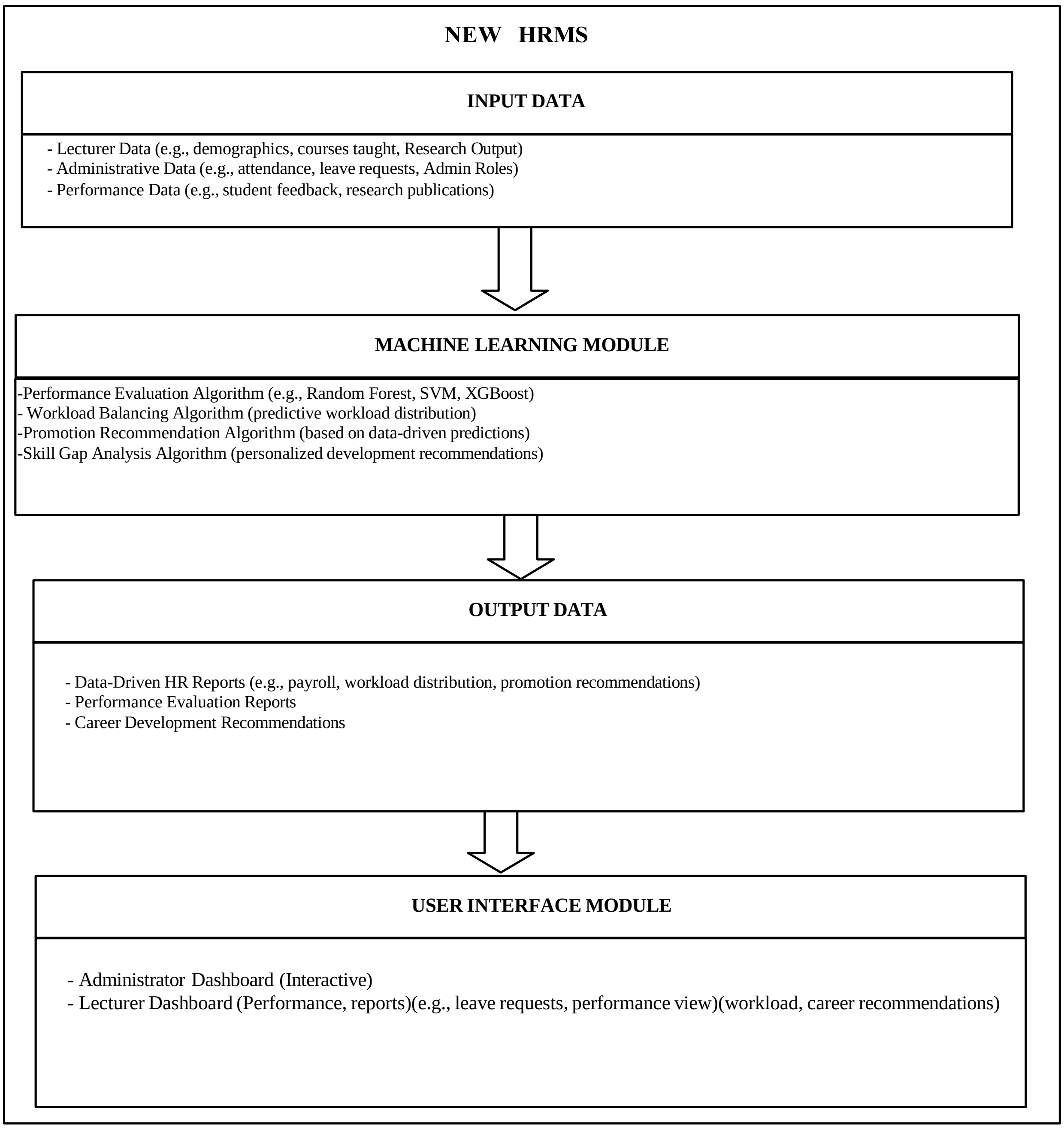


Table 1: Dataset Description

DOI: https://doi.org/10.5281/zenodo.17113277

Variable	Type	Description	Preprocessing
Teaching Load	Numeric	Hours per semester	Normalization, outlier removal
Research Output	Numeric	Number of publications, citations	Scaling, missing value imputation
Administrative Duties	Categorical	Roles: HOD, Dean, Committee member	One-hot encoding
Professional Development	Categorical	Training attended, workshops	Encoding, missing value handling
Years of Service	Numeric	Total years in university	Normalization
Promotion Status	Categorical	Eligible / Not Eligible	Encoding for classification



DOI: https://doi.org/10.5281/zenodo.17113277

4.1 Model Evaluation Metrics

The three Machine Learning (ML) models—Random Forest (RF), Support Vector Machines (SVM), and Extreme Gradient Boosting (XGBoost)—were evaluated using accuracy, precision, recall, and F1-score. These metrics balance predictive correctness with fairness in classification.

4.2 Comparative Performance of Algorithms

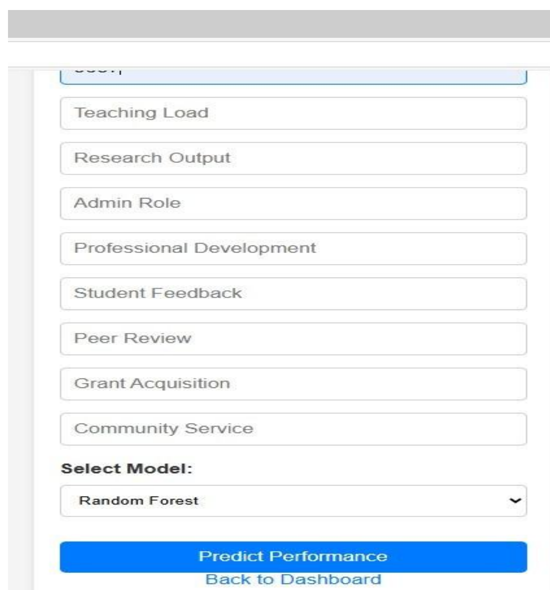
All models performed well, with different strengths. RF achieved over 85% accuracy and handled diverse features but was less interpretable. SVM was reliable in binary tasks like promotion prediction but weaker in multi-class settings. XGBoost consistently achieved the highest scores (above 90% accuracy, best F1), capturing complex interactions effectively.Table 2 summarizes the comparison of model performance.

Table 2: Comparative Performance of Machine Learning Models

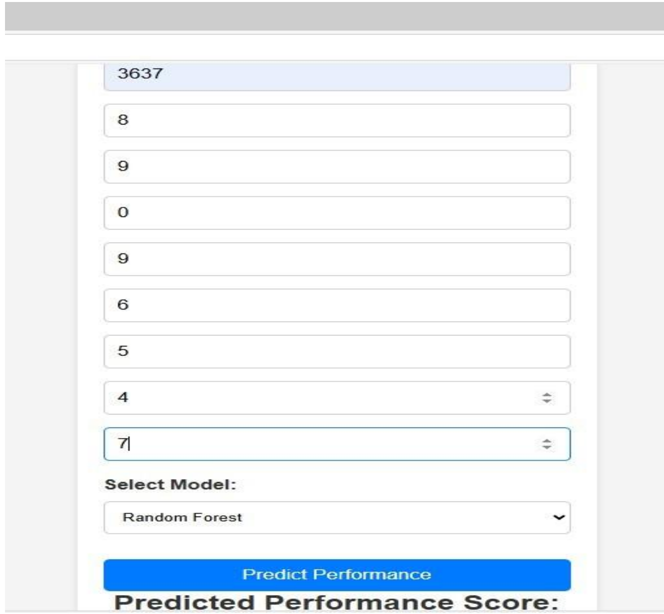
Metric	Definition	Random Forest	SVM	XGBoost
Accuracy	Percentage of correct predictions	89%	85%	92%
Precision	Correct positive predictions vs. all positives	88%	83%	91%
Recall (Sensitivity)	Correctly identified positives	86%	81%	90%
F1 Score	Harmonic mean of precision and recall	87%	82%	91%
Response Time (ms)	Time taken to process requests	120ms	150ms	110ms

4.3 System Prototype Outputs

A functional web-based prototype was developed to demonstrate the practical usability of the proposed HRMS. The system integrates the trained ML models into an intuitive interface that provides administrators and lecturers with real-time insights. The dashboard presents summaries of lecturer performance, workload distribution, and research-to-teaching ratios, enabling a clear overview of institutional activities. The promotion predictor delivers automated, data-driven recommendations while highlighting the key factors influencing eligibility. To support equitable task allocation, the workload balancer visualizes teaching and administrative responsibilities across departments, helping administrators redistribute tasks more fairly. Additionally, a skill gap analyzer identifies areas where staff may require professional development based on their performance metrics. Figures 3 and 4 illustrate the input interface and



DOI: https://doi.org/10.5281/zenodo.17113277



sample predictive outputs.

Figure 3: Lecturer's metrics input form

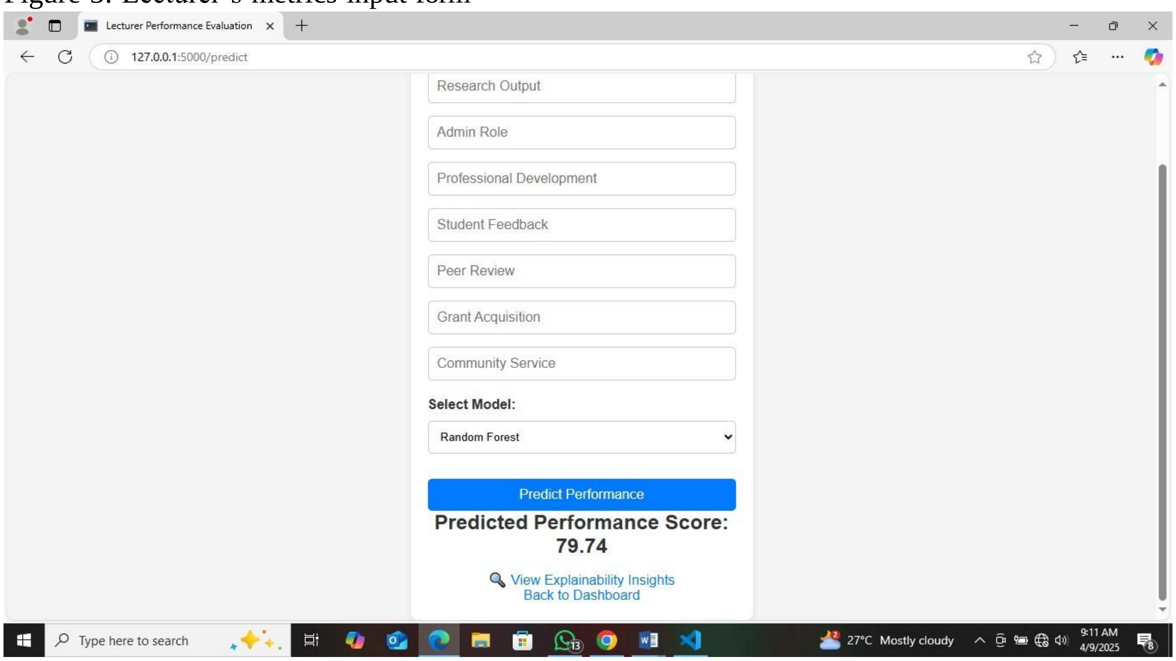


Figure 4: Predictive Outputs

4.4 Summary of Findings

The results indicate that XGBoost achieved the highest predictive accuracy, supported by strong F1-scores and robustness in handling complex relationships within HR data. Random Forest and SVM also performed well, offering complementary strengths in different contexts. When combined with the web-based HRMS prototype, these models provide a scalable and effective framework capable of reducing bias, optimizing workloads, and promoting transparency in staff promotion processes within Nigerian universities.

V Discussion

The findings confirm the effectiveness of machine learning in addressing critical HRM tasks such as workload balancing, performance evaluation, and promotion

DOI: https://doi.org/10.5281/zenodo.17113277

prediction. By embedding trained models within a modular, web-based framework, the study demonstrates how AI can modernize HR functions in resource-constrained universities and foster more transparent decision-making.
Among the models tested, XGBoost consistently achieved the highest predictive performance, owing to its ability to capture

complex, non-linear relationships in HR data. Random Forest offered stable and

generalizable results across tasks, while Support Vector Machines proved reliable in binary classifications but struggled in multi-class contexts. These insights highlight the promise of ensemble methods—particularly boosting techniques—as the most suitable approach for HR analytics in higher education.
The system's design further strengthens its relevance by ensuring scalability. Built with object-oriented principles and deployed using Flask and SQLAlchemy, the HRMS can be extended to incorporate additional functions such as recruitment analysis, attrition prediction, or training recommendations. Although tailored to Nigerian universities, its modular architecture enables deployment in similar institutional contexts across other developing regions facing comparable HR challenges.
Nevertheless, the study faced limitations. Data availability and quality posed significant challenges, as institutional records were often incomplete or inconsistent. While preprocessing mitigated some of these issues, larger and more reliable datasets would enhance model robustness. Another limitation lies in interpretability: models like XGBoost, while highly accurate, function as black boxes, which may hinder adoption by HR administrators. Future work should therefore integrate explainable AI methods and extend testing across multiple institutions to validate scalability and real-world applicability.

VI Conclusion And Future Work

This study developed a machine learning-driven HRMS based on the CRISP-DM framework to address biased evaluations, unbalanced workloads, and delayed promotions in Nigerian universities. The framework contributed in three key ways: it provided a comparative evaluation of Random Forest, SVM, and XGBoost, with XGBoost emerging as the most effective; it introduced a scalable, modular, web-based HRMS prototype; and it integrated predictive analytics into HR processes, thereby enhancing transparency, fairness, and efficiency.
Looking ahead, further research should explore deep learning approaches such as recurrent or graph neural networks to capture more complex patterns in HR data. Attention should also be given to explainability and fairness by integrating bias auditing and XAI techniques, ensuring that predictions are transparent and equitable. Large-scale validation across multiple universities would provide evidence of generalizability, while expanding the system to include functions such as recruitment, attrition analysis, and professional development tracking would broaden its utility.
In conclusion, the study shows that ML-powered HRMS solutions can serve as transformative tools in higher education, offering data-driven, transparent, and scalable pathways to improve staff management and institutional effectiveness.

References

- Afolabi, I. T., Olanrewaju, R. F., & Salami, A. O. (2019). Academic staff performance appraisal system using rule-based techniques. International Journal of Computer Applications, 182(35), 1–7. https://doi.org/10.5120/ijca2019918504
- Armstrong, M., & Taylor, S. (2020). Armstrong's handbook of human resource management practice (15th ed.). Kogan Page.
- Bock, L. (2015). Work rules!: Insights from inside Google that will transform how you live and lead. Twelve.
- Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. https://doi.org/10.1007/BF00994018
- Davenport, T. H., Harris, J., & Shapiro, J. (2010). Competing on talent analytics. Harvard Business Review. https://hbr.org/2010/10/competing-on-talent-analytics
- Guenole, N., Ferrar, J., & Feinzig, S. (2017). The power of people: Learn how successful organizations use workforce analytics to improve business performance. Pearson Education.
- Mahmud, H., Alam, M. R., & Hasan, M. (2018). A decision support system for promotion of faculty members. International Journal of Advanced Computer Science and Applications, 9(12), 123–129. https://doi.org/10.14569/IJACSA.2018.091216
- Nawaz, N., & Gomes, A. M. (2020). Artificial intelligence chatbots are new tools for customer satisfaction in the hospitality industry. International Journal of Advanced Computer Science and Applications, 11(1), 536–547. https://doi.org/10.14569/IJACSA.2020.0110168
- Oyelade, O. J., Ezugwu, A. E., & Oladipo, O. F. (2020). Machine learning model for academic staff appraisal in Nigerian universities. International Journal of Advanced Computer Science and Applications, 11(8), 45–53. https://doi.org/10.14569/IJACSA.2020.0110807
- Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 469–481). ACM. https://doi.org/10.1145/3351095.3372828
- Saint, W., Hartnett, T. A., & Strassner, E. (2003). Higher education in Nigeria: A status report. World Bank.
- Sharma, R., & Sharma, R. (2020). Artificial intelligence in human resource management: A conceptual perspective. International Journal of Management, 11(5), 1–10. https://doi.org/10.34218/IJM.11.5.2020.001
- Tursunbayeva, A., Franco, M., & Pagliari, C. (2018). Use of artificial intelligence in human resource management: A systematic review. Journal of Human Resource Management, 18(2), 20–36. https://doi.org/10.21511/imfi.18(2).2018.03
- Tursunbayeva, A., Di Lauro, S., & Pagliari, C. (2018). People analytics—A scoping review of conceptual boundaries and value propositions. International Journal of Information Management, 43, 224–247. https://doi.org/10.1016/j.ijinfomgt.2018.08.002
- Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining (pp. 29–39).
- Zhang, Y., Zhao, Y., & Xu, X. (2020). Machine learning approaches to employee performance evaluation: A comparative study. Journal of Computational and Applied Mathematics, 369, 112569. https://doi.org/10.1016/j.cam.2019.112569

DOI: https://doi.org/10.5281/zenodo.17113277