AI Based Decision Support Systems for Human Resource Management

Iman Thajeel Hillo and Kadhim Hasen Alibraheemi
Department of Computer Science and Artificial Intelligence, College of Education for Pure Sciences, University of Thi-Qar, 64001 Nasiriyah, Iraq
eman_thjeel@utq.edu.iq, Khalibraheemi@utq.edu.iq

Keywords: Absenteeism, CatBoost, Class Imbalance, Promotion, CatBoost, Machine Learning.

Abstract: Human Resour (HR) departments are leveraging data-driven decision-making for more effective management of absenteeism, promotions, and retirement planning. Conventional methods are often slow and subjective, especially in big companies like oil firms. In this research, we propose an AI-enhanced Human Resource Management System (AI-HRMS) which consolidates absenteeism forecasting, promotion prediction and retirement risk estimation into one decision-making assistance framework. Using actual HR data from the oil industry, we pre-process this raw data and build the physical world models using Random Forest, CatBoost and Artificial Neural Networks (ANN). SMOTE or under-sampling techniques were used to balance the datasets based on the type of prediction task. All models were evaluated with accuracy, precision, recall, F1-score, ROC-AUC and PR-AUC. The models are deployed as standalone AI services in a Docker container and seamlessly integrated into HRMS for real-time inference. Overall performance was excellent for the identification of absenteeism risk, prediction on retirement, and estimation of promotion eligibility for all models. The integration was used to provide real-time decision support to HR managers. After all, use of AI in supporting HR planning through the multi-model AI-HRMS system supports reduction of manual work and is able to maintain consistency in large organization decision making and help establish a modular architecture that is future-ready for integration with a host of HR analytics systems.

1 INTRODUCTION

Artificial intelligence (AI) is becoming a critical ingredient for the digital transformation of oil and gas operations as it complements human workforce in effectively managing complex processes and data-intensive tasks. State-of-the-art in AI, computational power and the analysis of big data, as well as machine learning now provide new tools to be implemented for decision making in exploration, production and reservoir management. That's important, as oil and gas companies feel the heat to streamline, mitigate risk and better use data to manage their fields [1]. Containerized-based architectures (such as Docker) are becoming increasingly important in terms of operating AI services in modern corporations. Docker allows us to containerize so that we have more convenient, reproducible environments that will be easier to manage and scale for those machine learning apps. A great fit for higher-end HR service delivery platforms that require efficiency and flexibility to manage sophisticated employee administration [2]. Staff promotion is a multi-dimensional problem in the HRM. Manual checking is prone to laborious, time-consuming and inconsistent, reducing the effectiveness of promoting. Hence, experts from the industry are relying more on machine-learning models for promotion approach to make it fairer, unbiased and based on reality [3]. Large-scale studies using advanced techniques such as Gradient Boosting and the k-Nearest Neighbors algorithm have repeatedly found that machine learning-based forecasting methods perform well in modeling and predicting promotions or other career-related outcomes in large public-sector organizations. These results highlight the important contribution and value added of HR analytics that are founded on data, in particular with regard to ensuring fairness and consistency in organizational procedures [4]. As well as the Sickness absenteeism-encompassing work absence based on a medical certificate (MC) due to health reasons-is an important organizational and economic issue in the workplace. Absence related costs constitute 1.2-2% of GDP in the OECD

countries to underscore its high impact. Considering that 1 in 5 working age people suffer from mental health problems, early intervention is crucial to avoid long-term absences. As such, the prediction of absenteeism based on data-driven AI methods provides a valuable tool towards minimizing disruptions and increasing workforce constancy [5]. Early retirement is thus an important breakpoint for both employees and organizations in that leaving the labor force earlier than average can significantly stress pension systems and erode long-term financial sustainability. Research has shown that a person's own characteristics, health and income as well as macroeconomic conditions all affect decisions to retire before age 65. The forecasting of them is important as early retirement result in less contribution and higher pension liabilities. Given that, machine-learning models can provide an effective method to recognize those at a higher risk and help plan better HR and pension systems [6]. This paper presents:

1) An AI-HRMS based containerized artificial intelligence system for promotion, absenteeism, and retirement prediction to Thi-Qar Oil Company. Utilizing Docker deployment.
2) Controlling and demonstrating the handling of imbalance HR data by using. The Synthetic Minority Oversampling Technique (SMOTE).
3) Contributing to digital transformation of oil landscape by delivery precise and expressive promotion recommendations which help in fair HR decisions.

Promotion, retirement, and absenteeism prediction are integrated in an HR decision-support framework in this study as these actions all rely on similar employee attributes including age, service years, job role, and performance. This consolidation eliminates data compartmentalization, empowering HR managers to access unified and consistent real-time predictive insights from a single platform.

1.1 The Synthetic Minority Oversampling Technique (SMOTE)

The Synthetic Minority Oversampling Technique (SMOTE) is an oversampling technique that copes with class imbalance issue by creating synthetic samples for the minority class instead of just replicating if there is one. It does so by generating realistic synthetic examples between a data point and its nearest neighbors, such that the classifier can better learn minority patterns. It leads to a cleaner, more balanced dataset, decreases class skewness and improves the machine learning model's performance [7] for example show in Figure 1.

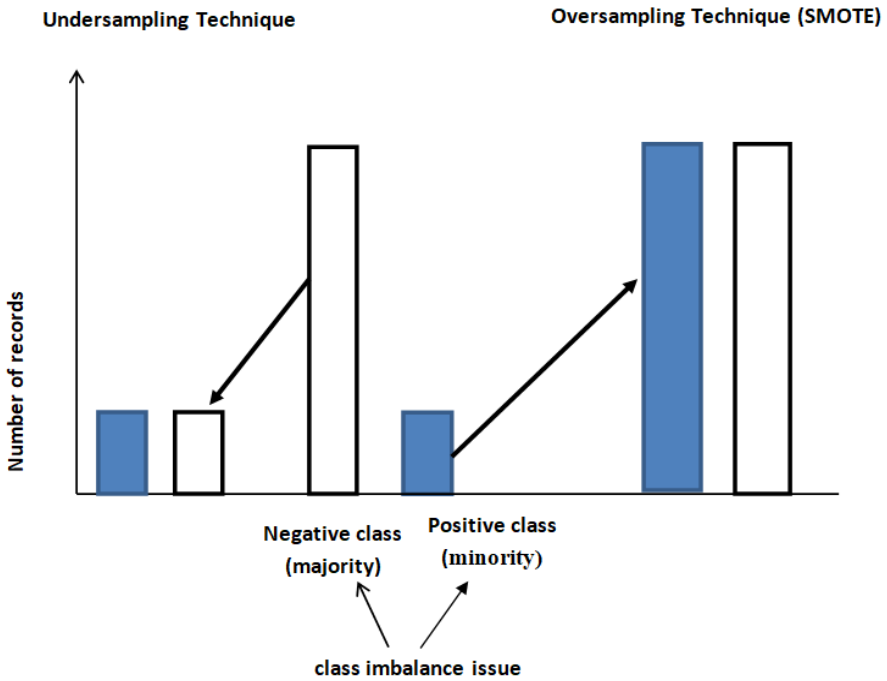


Figure 1: Imbalanced datasets.

1.2 Under Sampling Algorithm

Under-sampling is an approach for dealing with class imbalance, which equates sample sizes between the majority and minority classes. The algorithm filters out a portion of the majority-class pattern, then the learning classifier processes what remains, which almost lead to balance learning samples [3]. As in Figure 1.

2 RELATED WORKS

Resource Optimization with Intelligent System for Human Resource Management Recent research reflected on the implementation of intelligent system for HRM. Researches have disclosed critical elements in promotion, Retirement and Absenteeism decision-making from statistical models and machine learning to detect patterns hitherto overlooked by traditional approach. Also, the role of AI in decision-making processes via predictive algorithms and data augmentation has been discussed. Service-oriented designs to be reliable in complex environment. These investigations lay the foundation to build an AI-augmented HRMS which yields high accurate predictions and good operational performance. Mohammed and Mudhafar (2019) examined the HRIS to support HRM strategies via a study with 64 academicians while revealing that HRIS enhance recruitment, performance appraisal, decision quality; however this study is based on self-report data collected from small sample size and it contained no operational or system level analysis [8]. The early literature on HRMS concentrated in the analysis of classical multi-layer architectures (e.g., J2EE, MVC) that allow better organizational structure but are monolithic and hard to scale. Recent work by Kim et al., on the other hand. (2022) have shown the superiority of docker containerization as a way to obtain modular and light-weight deployment of machine-learning services, allowing more modular and runtime-driven systems. This trend away from monolith design and towards use of containers is an indication of the direction technology is heading and one element driving us to build our own AI capable HRMS [2]. Openja et al. (2022) investigated the usage of Docker for deployment in machine-learning software projects with a systematic study involving 406 GitHub repositories. Their approach relied on the manual annotation of project categories, to analyze Dockerfiles and collect Docker Hub images concerning their use. Through analysis, the study defined six ML project categories and 21 general uses of Docker that provide tangible insights for engineers deploying ML systems [9]. Bai et al. (2023) present a user-friendly AI environment by incorporating Dockerised JupyterLab environments into Galaxy for reproducible machine and deep learning workflows. Their approach encompasses: running Jupyter notebooks that leverage GPU acceleration, ONNX models, medical image segmentation using a Unet model, and automated dataset pipelines all executed inside containers. Experimental results showed high precision and recall, besides the dramatic reduction in training and inference time. The main contribution of this work is to make large scale, reproducible and user-friendly AI research possible by means of containerization [10]. Horodyski (2023) researched the attitudes of applicants towards AI recruitment applying TAM to evaluate perceived usefulness and ease of use. This was followed by a coinvestigator-administered quantitative survey, measuring applicants' expectations, perceived fairness and concerns about accuracy and bias. The results suggested a generally favorable opinion of AI tools, though participants also pointed out concerns like the lack of human judgment and possible algorithmic bias [11].

Fen and Rong Yong (2023) developed a cloud-based HRMS system based on the JSP, J2EE and MVC pattern Therein they integrated the Google App Engine with BigTable Datastore enabling scalable solution with low-cost standalone browser. Their system combines eight functional modules (including staff files, contracts, training, attendance, salary and performance management) as well as leverages distributed cloud storage for enhanced querying speed and reduced information technology maintenance loads. The work, however, was only at a conceptual level and it was not implemented nor tested in any enterprise system, hence there were no empirical performance results or study of security implications for such deployment methods on enterprise scale systems scalability and practicality [12]. Concomitantly, as literature on AI-driven predictive analytics in HR deepened; AI-based HR predicting analysis had also unfolded gradually, like Wang et al. (2024) discussed the state of artificial intelligence in oil and gas industry, especially its advantage for processing geology and production big data. The research demonstrates how AI is used for reservoir analysis, production forecasting, intelligent monitoring and operational optimization, allowing decision-makers to make more accurate decisions in data-heavy environments. While this is

not the focus in HR, their results show how AI may potentially revamp complex enterprise systems, which can provide lessons using a larger scale organization at digitalizing HR processes [1]. Căvescu and Popescu (2025) who conducted a systematic review on AI-driven HR predictive analytics, revealing the increasing involvement of machine learning algorithms e.g., XGboost, Random Forest, SVM and logistic regression in employee turnover prediction and workforce management. This, their findings reveal, can help to automate HR tasks, mitigate bias and enhance decision-making throughout recruitment, retention and performance optimization-particularly when it comes to skewed HR datasets [13]. Recent system-level studies also investigated AI incorporation in microservices architecture. (2023) conducted a systematic mapping study analyzing how AI methodologies are adopted in the various phases of microservices life-cycle. Their survey reveals that AI is applied in performance monitoring, anomaly detection, deployment optimization and operational automation. This does work shows increasing possibilities to include intelligent components into distributed architecture space, a fact which is also of immediate interest for today's AI enhanced HRMS tools [14]. Rocha Salazar and Boado-Penas (2019) developed a machine-learning based scoring system to predict early retirement out of individual characteristics and macroeconomic factors. They employed three supervised learning models-Random Forest, Logistic Regression, and Support Vector Machine-to predict workers retiring prior to 65. The logistic regression model achieved the highest AUC value (0.8277), suggesting that it is suitable for retirement risk scoring. The research suggests that pension providers can benefit from predictive modelling for avoid early retirements and making proper financial preparations [6]. Ajmi (2019) employed four methods using machine learning, Neural Networks, Decision Trees, Support Vector Machines and Logistic Regression to predict employee absenteeism in a BRAZILIAN courier company building the model on 721 observations and 21 attributes. The models were trained on a 70/30 split and tested using accuracy and AUC. The best performing model was Decision Tree with 83.33% accuracy and an AUC of 0.834, followed by the worst-performing SVM[15]. Shafie et al. (2023) solved the issue of unfair as well as slow decision making in promotion that is caused by imbalanced HR data, using Kaggle “HR Analytics: Employee Promotion” dataset which has only about 8.5% promoted staffs include 29853 rows. They performed two hybrid sampling methods - SMOTE+ENN and SMOTE+Tomek to balance the classes, combined with eight supervised learning models (Logistic Regression, Decision Tree, Random Forest, KNN, SVM, Naïve Bayes, AdaBoost and XGBoost) and three feature selection techniques (RFE-RFC, EVR-PCA, RFI-ECT). They utilized precision, recall and F1-score instead of the plain accuracy, to which XGBoost with SMOTE+ENN on a subset consisting only of eight important features (region, department, previous-year rating, KPIs >80%, awards won, age, recruitment channel and average training score) [16] . Ilwani et al. (2023) utilized three machine learning models (Logistic Regression, SVM and Random Forest) to Kaggle’s promotion dataset. Random Forest model performed the best by achieving 58.2\% for the recall, 100\% for the precision, 73.6\% for F-measure and 91.7\% -for accuracy. These findings suggest that RF discriminates effectively between promotable and non-promotable employees in this dataset [3]. In (2023) Chen, Lin, and Zhan introduced an employee promotion model in HR with a set of Kaggle dataset, the RandomForest classification and Analytic Hierarchy Process (AHP). After pre-processing the raw data, they compared RandomForest with logistic regression and apply AHP to rank employees for promotion. Tuned RandomForest resulted in ~93% accuracy and F1 = 0.94, yet the analysis was performed using only one public dataset, oriented on accuracy concentration without either time-aware validation nor real HRMS application [7]. Yasmine ,Aya Ibrir and Mahmut Çavur (2024) proposed a machine learning (ML) model in order to predict employee promotions based on an indicator of personal and professional attributes. A number of algorithms were tested and XGBoost (94 % accuracy, 94 % precision, 94 % recall and 94% weighted F-measure) gave the best results. The work showed good predictability but did not consider the class imbalance problem or real-system deployment, thereby leaving practical implementation and data-imbalance handling as critical gaps [17].

While previous papers mentioned HRMS design, investigated the AI and machine learning in predicting employee behavior, highlighted PR-AUC for imbalanced data as an important issue, these were disparate initiatives and disconnected. Precedent works treated the system design separated from predictive models, which were not combined into a functional HRMS.

The missing link, hence, is a contemporary HRMS built on microservices and Docker, integrated with AI in terms of promotion absenteeism

prediction, retirement prediction as well as handling the imbalanced data and deploying its predictions directly into the functional HR system.

3 METHODOLOGY

This paper describes the global methodology used to develop intelligent predictors implemented within the proposed AI-HRMS. The approach is structured in two main frameworks. The first framework is for both promotion and retirement prediction since two tasks are using same demographic and job- related traits of staff members as well as same decision properties. The second one is designed specifically for absenteeism's prediction given the differences in data structure, behavioural dynamics and factors of influence when it comes to employee attendance. All records for employees were anonymized and processed according to privacy and confidentiality guidelines.

3.1 Methodology for Predicting Promotion and Retirement

The retirement prediction task shares the exact same preprocessing methods and methodological scheme used for the promotion prediction model in order to provide fair and consistent evaluation across HR analytics tasks. In the common pipeline we have to clean the data and deal with missing & outlier values, normalize numerical features by applying Min-Max scaling, transform categorical variable using One-Hot Encoding. The class imbalance problem is dealt with SMOTE resampling method, while feature selection is executed by Point-Biserial Correlation Coefficient to keep highly influential variables. Next, the five supervised learning classifiers (Logistic Regression; Random Forest; XGBoost; LightGBM and CatBoost) are implemented and tested under the same thrK conditions. This standardized approach allows the promotion and retirement prediction models to be tested consistently when implementing within AI-HRMS. The proposed methodology for Promotion and Retirement is shown in Figure 2.

3.1.1 Dataset Explain

Promotion and retirement datasets contains (2400 record for promotion and 2400 for retirement) These are real-world data obtained from information system of Thi-Qar Oil Company human resources management. Each record corresponds to an employee and contains demographic and work-related features (age, service length, job title, department, level of education,…), the salary other: training records and appraisals. Both data sets have the distinct problem of imbalanced class, on which most employees are non-promoted and non- retiring. This imbalance is evident in Figure 3 for promotion and retirement and was the impetus to consider appropriate imbalance-handling techniques during model-building.

3.1.2 Preprocessing

The dataset is first separated into numerical and categorical features. Whenever data is missing for a certain subject, the average (mean for numerical variables and mode for categorical ones) of this particular population is included. Numerical outliers, more than three standard deviations above or below the mean, are replaced by their corresponding feature mean to reduce the effect of extreme values.

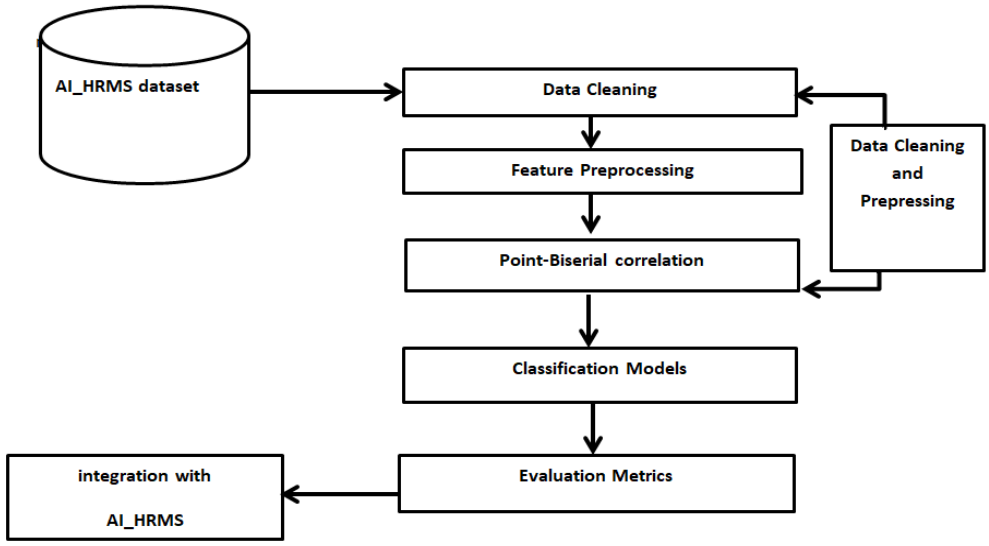
Following data cleaning, all numerical features are scaled to the range [0,1] according to Min-Max normalization. The categorical features are converted by one-hot encoding, which produces binary indicator variables suitable to ML methods.

The problem of class imbalance was added as an add-on stage after pre-processing was done. Synthetic Minority Oversampling Technique (SMOTE) and under sampling methods were tested. These methods were found to have the most stable result and was able to maintain the structure of minority class while enhancing learning process for models.

3.1.3 Point-Biserial Correlation Analysis

Point-Biserial correlation analysis is performed to reduce dimensionality and keep only statistically significant predictors. Non-significant features ($p \geq 0.05$) are removed and correlations of the reduced representation feature set, judged by the magnitude of their absolute correlation with response values, ranked. When training, we choose the top 40 features for training in order to improve performance while avoid overfitting. Table 1 Point-biserial correlation between the point inputs.

The Table 1 gives the correlation coefficient, statistical significance (p), and absolute value of a level's correlation with T before training the model. These rankings aid in the identification of which variables are most strongly related to promotion decision.



Figuer2: The proposed methodology for promotion and retirement.

Figure 3: The class imbalance of the AI-HRMS dataset.

Table 1: Point-biserial correlation between the point inputs (promotion).

Feature	Correlation Coefficient	P-value (Statistical Significance)	Absolute Correlation Coefficient
Thanks_Months_48m	-0.180392	6.89584E-35	0.180392
Effective_Service_Months	0.119829	3.72236E-16	0.119829
Months_Since_Last_Promotion	0.10742	2.91246E-13	0.10742
Penalty_Months_Total	0.104393	1.32567E-12	0.104393
Eligible_Now_48m	0.086957	3.58841E-09	0.086957

Table 2: Point-biserial correlation between the point inputs (retirement).

Feature	Correlation Coefficient	P-value (Statistical Significance)	Absolute Correlation Coefficient
Age_Years	0.528443	9.34E-173	0.528443
Service_Years	0.043926	3.14E-02	0.043926
Y_RETIRE_WITHIN_6M	0.693195	< 1.00E-300	0.693195
Y_ALREADY_ELIGIBLE	−0.062704	2.12E-03	0.062704

We can see from the results in Table 2, retirement prediction is mostly affected by age agd service attributes. Age_Years and previous retirement signals are the most important features for the target, in line with known deterministic and regulatory nature of retirement intentions. Other variables, as department affiliation, present just local contributions and perform the part of secondary variations to age and service duration.

3.1.4 The Models Used

Five supervised classifiers are implemented using the pre-processed and balanced dataset:

- (LR) with more number of iterations (max_iter = 2000).
- (RF) with 300 Estimators (n_estimators = 300).
- XGB with tuned boosting parameters.
- LGBM with Histogram-Based Gradient Boosting.
- CatBoost (CB) with categorical-feature-induced boosting.

The dataset is split into 70% as training and 30% as testing, stratified sampling is employed to ensure uniform class distribution. The same feature, hyper parameter and experimental setting is applied in the training of all models to have a fair performance comparison.

3.1.5 Evaluation of Models Performance

To comprehensively evaluate the effectiveness of the proposed promotion prediction models, several widely used classification metrics were employed, including Accuracy, Precision, Recall, F1-score, ROC-AUC, and PR-AUC [35–37]. These metrics are commonly used in machine learning research and provide complementary perspectives on classification performance.

Given the significant class imbalance in the promotion dataset, relying solely on Accuracy could lead to misleading conclusions. Therefore, greater emphasis was placed on Recall, F1-score, and PR-AUC, which provide a more informative assessment of the model's ability to identify employees eligible for promotion. Recall evaluates the proportion of actual promotion cases correctly identified by the model, while Precision measures the reliability of positive promotion predictions. The F1-score balances these two measures and is particularly useful when both false positives and false negatives are important.

In addition, ROC-AUC was used to assess the discriminative capability of the models across different decision thresholds. Since ROC-AUC may provide overly optimistic estimates in highly imbalanced datasets, PR-AUC was also considered because it focuses on the minority class and better reflects performance in promotion prediction tasks.

Together, these evaluation metrics provide a comprehensive framework for assessing both overall predictive performance and the ability of the proposed models to accurately identify employees eligible for promotion.

3.2 Methodology for Predicting Absenteeism

The absenteeism data was then loaded and the target absence label created. The data was preprocessed by feature scaling and categorical encoding. In each fold of the 5-fold cross-validation, 80% samples were used for training, while the remaining samples were left to evaluate performance. SMOTE was performed only on the training data to solve class imbalance problem. For each fold, four machine learning classifiers including Logistic Regression, Support Vector Machine (SVM), Random Forest, and Artificial Neural Network (ANN) were fitted and tested. Model evaluation was based on the Accuracy, F1-score and AUC-ROC measures with its average over the five folds as final results.

3.2.1 Absenteeism Data Upload

The absenteeism database in this work contains 25,600 data records with 32 attributes taken from the HR database. The dataset contains the information about past attendance in hours and several demographic, job-related and behavioral details of employees. The target is the employee's absence status in a time period. The data set reflects the reality of company operations and there is a natural asymmetry between absentee and non-absentee cases, which positions it to be useful for developing and testing models. Class distribution show in Figure 4.

3.2.2 Data Preprocessing

Before the absenteeism dataset was ready for machine learning modelling, it underwent preprocessing. 1 All input features were examined and separated into numerical and categorical features. The numerical features were normalized by feature scaling, to ensure all the variables bound to the comparable range and prevent large magnitude features from weighting more on learning. Categorical variables were encoded into numeric forms for machine learning models. The preprocessing pipeline was trained on the training

data and used to transform test data within each fold, in order to avoid leaking information from the test set. This pre-processing step allows to standardize the scale of the input features and facilitates model convergence thus improving overall predictive performance.

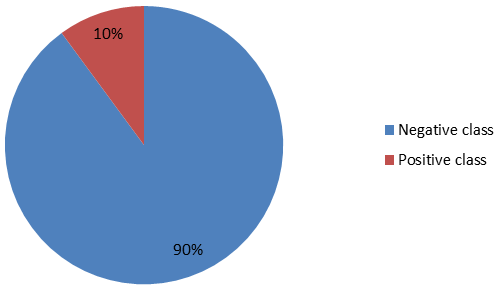


Figure 4: The class imbalance of the absenteeism dataset.

3.2.3 Cross-Validation and Data Splitting

For model evaluation, stratified 5-fold cross-validation was used to keep a reliable measurement of performance and minimize sampling bias while keeping the original class distribution in each fold. This method takes the dataset and splits it in five equal subsets (folds) using these folds, we train on 4 of them, then test on remining subset. This procedure was repeated five times, with each of the folds used as a test set once. Stratification ensures that both training and testing set have representations of absentee and non-absentee classes in proportion, which is essential for imbalanced datasets. This approach improves robustness of the estimation and minimizes the chance of overfitting and performance bias.

3.2.4 Handling Class Imbalance

To mitigate the problem with class imbalance in the absenteeism data, we used SMOTE, which transformed only the training examples in each fold of cross-validation. In contrast to this, SMOTE creates new samples for the minority class by considering its feature space similarity instead of merely replicating them.

3.2.5 Model Training and Evaluation

The four machine learning approaches, Logistic Regression, SVM, Random Forest and ANN were developed in each fold of the cross-validation scheme. Model performance was evaluated by Accuracy and F1-score as well as the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) that quantifies overall correctness, class discrimination, and predictive power. The average of all test performances over the 5-fold runs was taken as the final performance of each model, thus guaranteeing a reliable and unbiased comparison among models.

3.3 Containerization and Operationalization of the Models

To operationalize the machine-learning models, a lightweight deployment workflow was implemented with Docker. The promotion, retirement and absenteeism prediction models were encapsulated into separate docker containers adjacent to their respective Python environment and the required dependencies. These model containers communicate with a separate API container over REST endpoints and the database runs in an another isolated container. So this containerized architecture ensures the reproducibility, easy management of environment and also prediction services can be deployed on any machine with minimal configuration. Docker Compose is used to orchestrate all the containers bringing together the system into a running integrated scalable HR-decision support platform.

4 RESULT AND DISCUSSION

In this section we discuss our experimental results for the promotion, retirement and absenteeism prediction tasks in two scenarios without data balancing, and with an approach that balances the data using SMOTE to handle class imbalance. Various kinds of machine learning classifiers were compared using the performance measures of classifier and those with optimal performance measures who applied to build the AI-HRMS.

4.1 Analysis of Promotion Results

Though there are no measures as precision, recall and f1-score for promotion being zero because of heavy class inbalance due to No Promotion, but classification is accurate according to greater deluge with accuracy $\approx > 0.956$ XGBoost and LightGBM.Logf= xgboost → LightGBM achieved considerable accuracy (≈ 0.956), i.e there was failure to detect the positive class-Positive promotion{2} induced by severe imbalance of classes. CatBoost showed better Precision, Recall and F1-score as well compared to the others, which means it does have

some capability in dealing with the imbalance, but still performs much worse than balanced scenario. The results are presented in Table 4.

All methods showed impressive performance improvement after class imbalance problem was resolved by under sampling method. CatBoost performed best out of the models tested because its approach to categorical variables is better, it uses ordered boosting which reduces overfitting, and it excels with small-to-medium structured datasets. This characteristic resulted in its relatively higher Accuracy, F1-score, ROC-AUC and PR-AUC displayed in Table 5. These findings suggest that data balancing is indispensable for trustworthy promotion prediction models of AI-HRMS.

4.2 Analysis of Retirement Results

Pre-balance (unbalanced learning) outcomes reveal that Logistic Regression has very low recall (0.7037) although high accuracy and ROC-AUC scores, suggesting a bias towards the dominant class with less identification of true retirement cases. Even though Random Forest shows good results in all the metrics, the imbalanced dataset may have an impact on its results. Smoothing skills are required to increase the detection of positive class. Table 6 presents the retirement prediction results before data balancing.

Table 4: Promotion prediction results before data balancing.

models	Accuracy	Precision	Recall	F1	ROC-AUC	PR-AUC
XGBoost	0.956	0	0	0	0.681	0.075
LightGBM	0.956	0	0	0	0.681	0.075
CatBoost	0.979	0.863	0.612	0.716	0.994	0.872

Table 5: Promotion prediction results after data balancing (SMOTE).

models	Accuracy	Precision	Recall	F1	ROC-AUC	PR-AUC
XGBoost	0.714286	0.782609	0.580645	0.666667	0.753024	0.670673
LightGBM	0.587302	0.555556	0.806452	0.657895	0.608871	0.553292
CatBoost	0.888889	0.928571	0.83871	0.881356	0.973286	0.970691

Table 6: Retirement prediction results before data balancing.

Model	Accuracy	Precision	Recall	F1	ROC-AUC	PR-AUC
Logistic Regression	0.976	0.974	0.703	0.817	0.997	0.968
Random Forest	1	1	1	1	1	1

Table 7: Retirement prediction results after data balancing (SMOTE).

Model	Accuracy	Precision	Recall	F1	ROC-AUC	PR-AUC
Logistic Regression+SMOT	0.991	0.983	1	0.991	0.999	0.999
Random Forest+SMOT	1	1	1	1	1	1

Table 8: Absenteeism prediction results before data balancing.

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
SVM+SMOT	0.901	0.8885	0.807	0.956	0.758
Random Forest+SMOT	0.991	0.9955	0.916	0.956	0.708
Logistic Regression+SMOT	0.8	0.69	0.700	0.859	0.654
ANN+SMOT	0.912	0.901	0.887	0.854	0.908

Table 9: Absenteeism prediction results after data balancing (SMOTE).

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
SVM+SMOT	0.991	0.948	0.967	0.956	0.998
Random Forest+SMOT	0.991	0.995	0.916	0.956	0.998
Logistic Regression+SMOT	0.968	0.78	0.970	0.859	0.994
ANN+SMOT	0.9948	0.9873	0.9652	0.9741	0.9988

Results of the post balancing show an enhanced behaviour of the model. That Logistic Regression has a higher F1 and recall compared to pre-balancing shows a better detection of retirement cases. Random Forest also attained 1 in all metrics (Accuracy = 1, Precision = 1, Recall = 1), which confirmed its stability for balanced data as shown in Table 7. This was a crucial step for improving the reliability of our model and minimizing biases.

This near-perfect performance is obviously due to one simple reason: the Random Forest model is effectively learning some specific, (almost) perfectly decidable rule among the retirement data. Since retirement is a phenomenon determined by a clear set of explicit conditions, the model excellently reproduces this legal-like decision boundary with high precision and as it becomes perfect on scores across all what we use for performance evaluation.

4.2 Analysis of Absenteeism Results

Performance of pre-balancing absence results These below in Table 8 pre-balancing absence result indicate performance variation caused by the class imbalance. The accuracy (0.9916) &AP=precision(0.9955) for Random Forest remains the highest; however, it has the lowest ROC-AUC suggesting tertiate to majority class sensitivity and a good performance on weighted reciprocal of precision recall plot compared to SVM. SVM achieved similar performance with smaller ROC-AUC (0.7584), and learn it hard on negative cases. Logistic Regression performed the worst on Precision and ROC-AUC, mainly perfomed poorly in imbalanced data. The ANN model performed with good balance and highROC-AUC(0.9088),which was more superior in discrimination behaviors. In summary, the results suggest that a trade-off exists between fairness of estimation and reliability.

Table 9 presents the absenteeism prediction results after data balancing using SMOTE. All models demonstrated substantial improvements in discriminative performance, with ROC-AUC values approaching 1.0, highlighting the importance of addressing class imbalance. SVM achieved a ROC-AUC of 0.998, while Random Forest maintained high precision and F1-score and showed a considerable improvement in ROC-AUC compared with the pre-balancing results. Logistic Regression exhibited a marked increase in recall (0.971) and ROC-AUC (0.994), indicating a substantially improved ability to identify absenteeism cases. The ANN model achieved the best overall performance in terms of accuracy, F1-score, and ROC-AUC after data balancing. Overall, these results demonstrate that data balancing techniques are essential for improving predictive performance and reducing bias toward the majority class.

Final models for each of the prediction task in the system were derived from the postbalacing evaluation performances. We selected CatBoost with under sampling as the end model for promotion prediction due to its highest performance in overall, especially accuracy, F1-score, ROC-AUC and PR-AUC on average across all metrics indicating strong discriminative power and generalization. To predict absence, the best model used was ANN due to its very good balanced performance after balancing which are justified by an accuracy, F1 and ROC-AUC high, whereas Random Forest stayed as a secondary model for their better robustness. On the prediction of retirement, Random Forest was chosen as the final model since after balancing, it obtained perfect performance across all evaluation metrics which can be credited to the fact it is able to learn well defined law-based rules behind retirement.

5 CONCLUSIONS

This research comprehensively demonstrates that the AI-HRMS effectively assists with three basic human resource tasks: promotion, retirement, and absenteeism prediction, utilizing highly reliable machine-learning models. The system uses advanced techniques, which lead to a highly accurate and objective decision-making process that is not founded on biased opinions or preconceived notions about the data, tailored specifically for HR functions these could include metrics such as age of employees, years of service within the company among others. In addition, it effectively captures the complex patterns of employee absenteeism, offering insights that can be used for planning purposes. Indeed, these striking outcomes speak volume on the viability of the system for organizations and how it could become one of their most useful tools in the scope of Human Resources. They also serve as a solid baseline for potential future developments in HR analytics and decision-support software design, which could lead to more innovative systems that transform how we manage our most important resource.

6 LIMITATIONS AND FUTURE WORK

This study has several limitations. The data were from one institution, which could impact generalizability and introduce bias. Some models

performed nearly perfectly (e.g. retirement model) as the data was rule-based, but external validation is lacking so robustness confirmation is reduced. Balancing methods (SMOTE and undersampling) may introduce synthetic noise or lose majority-class information, and none of the alternative techniques were statistically tested. There may be some inconsistency due to basic manual hyperparameter tuning and differences in validation strategies across tasks. There's also no wide interpretability or fairness analysis, which is key for deciding HR with AI. These issues will be addressed in future work by incorporating external datasets, use of strict validation practices for example nested cross-validation, and benchmarking against scalability and performance metrics. It will further study the explainability measures, fairness measures and comparative studies on balancing methods for enhancing the interpretability, robustness and applicability of the AI-HRMS system.

REFERENCES

[1] T. Wang et al., "Current Status and Prospects of Artificial Intelligence Technology Application in Oil and Gas Field Development," ACS Omega, Jan. 2024, [Online]. Available: https://doi.org/10.1021/acsomega.3c09229.

[2] B. S. Kim, S. H. Lee, Y. R. Lee, Y. H. Park, and J. Jeong, "Design and Implementation of Cloud Docker Application Architecture Based on Machine Learning in Container Management for Smart Manufacturing," 2022, [Online]. Available: https://doi.org/10.3390/app12136737.

[3] Y. Chen, X. Lin, and K. Zhan, "The Employee Promotion Decision based on the Randomforest Algorithm and the Analytic Hierarchy Process," 2023, pp. 1644-1653, [Online]. Available: https://doi.org/10.2991/978-2-38476-126-5_185.

[4] M. Ilwani, G. Nassreddine, and J. Younis, "Machine Learning Application on Employee Promotion," Mesopotamian J. Comput. Sci., vol. 2023, pp. 100-114, Jun. 2023, [Online]. Available: https://doi.org/10.58496/MJCSC/2023/013.

[5] H. Gülten and H. Baraçlı, "A Machine Learning-Based Forecast Model for Career Planning in Human Resource Management: A Case Study of the Turkish Post Corporation," Appl. Sci., vol. 14, no. 15, p. 6679, Jul. 2024, [Online]. Available: https://doi.org/10.3390/app14156679.

[6] N. Lawrance, G. Petrides, and M.-A. Guerry, "Predicting employee absenteeism for cost effective interventions," Decis. Support Syst., vol. 147, p. 113539, Aug. 2021, [Online]. Available: https://doi.org/10.1016/j.dss.2021.113539.

[7] J. de J. Rocha Salazar and M. del C. Boado-Penas, "Scoring and prediction of early retirement using machine learning techniques: application to private pension plans," An. del Inst. Actuar. Españoles, no. 25, pp. 119-145, Dec. 2019, [Online]. Available: https://doi.org/10.26360/2019_6.

[8] M. Altalhan, A. Algarni, and M. Turki-Hadj Alouane, "Imbalanced Data Problem in Machine Learning: A Review," IEEE Access, vol. 13, pp. 13686-13699, 2025, [Online]. Available: https://doi.org/10.1109/ACCESS.2025.3531662.

[9] M. A. Ahmad and M. H. Ali, "The Role of the Human Resources Information System in the Practice of Human Resources Management Strategies: A survey of the views of a sample of teaching staff at the faculties of Cihan University-Erbil," Acad. J. Nawroz Univ., vol. 8, no. 3, p. 83, Aug. 2019, [Online]. Available: https://doi.org/10.25007/ajnu.v8n3a409.

[10] M. Openja, F. Majidi, F. Khomh, B. Chembakottu, and H. Li, "Studying the Practices of Deploying Machine Learning Projects on Docker," 2022, [Online]. Available: https://doi.org/10.1145/3530019.3530039.

[11] A. Kumar, G. Cuccuru, B. Grüning, and R. Backofen, "An accessible infrastructure for artificial intelligence using a Docker-based JupyterLab in Galaxy," Gigascience, vol. 12, Dec. 2022, [Online]. Available: https://doi.org/10.1093/gigascience/giad028.

[12] P. Horodyski, "Applicants' perception of artificial intelligence in the recruitment process," Comput. Hum. Behav. Reports, vol. 11, p. 100303, Aug. 2023, [Online]. Available: https://doi.org/10.1016/j.chbr.2023.100303.

[13] "Design of Human Resource Management System Based on Cloud Platform," Adv. Comput. Signals Syst., vol. 7, no. 7, 2023, [Online]. Available: https://doi.org/10.23977/acss.2023.070707.

[14] A. M. Căvescu and N. Popescu, "Predictive Analytics in Human Resources Management: Evaluating AIHR's Role in Talent Retention," AppliedMath, vol. 5, no. 3, p. 99, Aug. 2025, [Online]. Available: https://doi.org/10.3390/appliedmath5030099.

[15] S. Q. Ajmi, "Predicting Absenteeism at Work Using Machine Learning Algorithms," Muthanna J. Pure Sci., vol. 7, no. 1, pp. 1-12, Dec. 2019, [Online]. Available: https://doi.org/10.52113/2/07.01.2020/1-12.

[16] S. Shafie, S. P. Ooi, and K. W. Khaw, "Prediction of Employee Promotion Using Hybrid Sampling Method with Machine Learning Architecture," Malaysian J. Comput., vol. 8, no. 1, pp. 1264-1286, Apr. 2023, [Online]. Available: https://doi.org/10.24191/mjoc.v8i1.18456.

[17] Y. A. Ibrir and M. Çavur, "Forecasting Employees' Promotion Based on Personal Indicators by Using a Machine Learning Algorithm," International Journal of Management Information Systems and Information Science, vol. 8, no. 2, pp. 75-98, Dec. 2024, [Online]. Available: https://doi.org/10.33461/uybisbbd.1471499.