Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 14, Issue. 6, June 2025, pg.85 – 99

PREDICTING RETIREMENT MOTIVES USING MACHINE LEARNING TECHNIQUES IN THE ZIMBABWEAN PENSIONS SECTOR

Rangarirai Dewah[1]; S Chaputsira[2]; Wilton Muzava[3]

[1]Data Science & Analytics & Harare Institute of Technology, Zimbabwe
[2]Computer Science & Harare Institute of Technology, Zimbabwe
[3]Computer Science & Harare Institute of Technology, Zimbabwe
[1] rangariraidewah@gmail.com; [2] schaputsira@hit.ac.zw; [3] wmuzava@hit.ac.zw

DOI: https://doi.org/10.47760/ijcsmc.2025.v14i06.010

Abstract: Zimbabwean pension fund stakeholders face challenges in managing retirement schemes and planning policies due to the limitations of traditional statistical methods in predicting retirement motives. These limitations contribute to mismatches in retirement timing, misaligned policies, and inadequate preparedness. Accurate, data-driven prediction models offer a promising opportunity to enhance retirement planning and policy alignment. This study investigates the application of supervised Machine Learning (ML) techniques to predict retirement motives among members of Zimbabwe's pension sector. Understanding retirement behavior is especially critical in emerging economies to design inclusive and sustainable social security systems. Using pension member data, the study evaluates the performance of various ML models, including Logistic Regression (LogReg), Support Vector Machine (SVM), Random Forest (RF), Multilayer Perceptron (MLP), and Extreme Gradient Boosting (XGBoost). Model performance was assessed using cross-validation with metrics such as accuracy, precision, recall, and ROC-AUC. SHapley Additive exPlanations (SHAP) were employed to interpret model outputs and identify key

features influencing retirement decisions—such as age, health status, job embeddedness, service length, financial literacy, and family responsibilities. Results show that ensemble models, particularly XGBoost and Random Forest, outperformed neural networks and linear models in predicting early and delayed retirement motives. The study highlights the potential of socially-informed ML approaches in low-resource contexts, offering valuable insights for policymakers and pension administrators. By integrating AI into pension forecasting, the research supports smarter fund management and policy development, contributing to more resilient and responsive pension systems.

Keywords: Machine Learning, Deep Learning, Logistic Regression, SVM, Random Forest, XGBoost ROC-AUC, SHAP, Heatmap, PCA, VIF, SMOTE, AI

I. INTRODUCTION

The Zimbabwean pensions sector, like many others globally, faces challenges in meeting the diverse needs of its members as they approach definitive abandonment of working life. Understanding the motives behind retirement decisions is crucial for pension fund administrators to tailor their services, enhance member satisfaction, and ensure financial sustainability. Traditional methods of analyzing retirement motives often rely on simplistic models that may not capture the complexity of individual decision-making processes. ML techniques, with their ability to handle large datasets and identify non-linear relationships, offer a promising approach to predicting retirement motives more accurately.

This study investigated the potential of ML in predicting why individuals choose not to retire, considering the unique socio-economic context of Zimbabwe. The paper makes significant contributions to academia and practice in the retirement analytics domain. It confirms that ML predictive models have the capacity to classify retirement motives and offer insights that outperform traditional statistical methods. The results demonstrate the potential of predictive modelling techniques particularly ensemble tree-based models in improving retirement analytics. By comparing five ML models the study identified the most robust and most effective models for predicting retirement motives using appropriate metrics. It identified key factors influencing retirement decisions in the Zimbabwean pensions sector. By accurately identifying the factors that influence retirement intentions, policymakers, and pension fund stakeholders, can design targeted interventions to increase retirement uptake and improve financial security. Overall, the paper provides valuable insights for promoting and optimizing retirement analytics to ensure a more sustainable retirement system.

II. REVIEW OF RELATED LITERATURE

A. Determinants of retirement motives

There is an extensive literature focused on understanding retirement motives rooted in literature domains covered in this study. According to [1], retirement refers to the transition from employment to joblessness. The traditional meaning of retirement is that someone works for many years and on reaching retirement age they withdraw from work completely [2]. However [3]'s study reveals that many retire before reaching retirement age through employee attrition. Both of these approaches are at variance with [4] and [5]whose findings are that come retirement age, would be retirees prefer to keep on working instead of exiting from work.

[6] assets that technology is key to tackling these retirement decision challenges yet most legacy pension providers in today's retirement industry run old legacy technology whose driving technique is record keeping suggesting its limitations as a system of record. They further posit that the future of retirement technology lies in techniques for security, data analytics, automated testing and ML. The Zimbabwean Insurance and Pensions regulator, the Insurance and Pensions Commission (IPEC) has added their voice reinforcing the need for using Artificial Intelligence (AI) techniques in the Zimbabwean Pensions Sector. According to the regulator of the

Insurance and Pensions Sector, IPEC, the Insurance and Pensions Industry needs appropriate technological innovation to generate actionable insights for insurance and pensions policy formulation".

[7] reports that "IPEC recognizes the importance of artificial intelligence in the insurance industry, given its potential benefits to the insurance value chain". This is supported by the Government adopted Vision 2030 one of whose tenets involves the introduction of a comprehensive e-Government system [8]. In Zimbabwe, a country with a unique socio-economic landscape, understanding retirement antecedents is crucial for individuals, employers, regulators, the government, NGOs (social workers) and pension service providers alike. This study seeks a supervised learning, classification examination of retirement motives using ML for both the early retirement and the delayed retirement motives.

According to [9], traditional approaches to understanding retirement decisions often rely on statistical analysis of personal, financial and demographic studies conducted using legacy technologies, which may not capture the full spectrum and depth of the influencing factors. However, [10], found that advancements in ML, offer new opportunities vis-à-vis traditional statistical techniques in data driven techniques for the prediction of retirement outcomes as ML models possess the capacity to analyze large datasets and uncover hidden patterns that influence retirement behavior.

These findings are supported by [11] who found that ML has been found to provide insights that are unattainable through traditional parametric methods and is more flexible for handling multi-dimensional data. According to [12], utilizing ML models in predicting retirement decisions is considerably new and is not yet widely documented in literature. He is supported by [13] who posits that ML has the potential to contribute much to the understanding of the retirement transition, the factors influencing it and its consequences. She further contents that there is need to go beyond demographic, health and wealth factors to understand the retirement process at the individual level.

Extant literature confirms [13] 's finding with [10] arguing that the issue of retirement prediction using ML techniques has so far been extensively applied in the economic and sociological domains mainly. According to [12], when carrying out retirement predictions, ML models utilize a multiple factors, a plethora of antecedents such as income level, pension savings, health status, age to name but a few, that influence retirement decisions.

This project builds on those findings and techniques and explores their applicability in the retirement domain to predict push factors that include uncharted motivators such as job satisfaction, family obligations, personal aggrandizement , continuation of service to name but a few. Some of the studies perused by this researcher so far reveal the utilization of ML models in early retirement predictions.

Logistic Regression (LR) and Decision Trees (DT) have been applied to model binary classification outcomes (e.g., whether an individual retires for financial reasons). For instance,[14] and [15] applied logistic regression model to predict early retirement motives based on age, income, and pension balances, achieving decent accuracy prediction on the financial motives. The application of Neural Networks for retirement motive prediction, particularly Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models, improved the prediction of dynamic retirement motives. [16] utilized LSTMs to predict early retirement motives based on longitudinal health and financial data, achieving a higher accuracy than traditional models. These models succeeded in capturing temporal dependencies, such as fluctuations in health and job working conditions over time.

The study conceptualizes retirement as a decision influenced by multiple push and pull factors, categorized into voluntary retirement, health-related exits, financial constraints, and policy-driven retirements [17]. These categories align with past research identifying income levels, pension savings, job satisfaction, health status, and workplace conditions as key determinants ([18]. The ML-based approach in this study aims to enhance traditional methods by integrating a wider range of features and improving prediction accuracy. The conceptual framework also incorporates government policies such as Zimbabwe's Vision 2030, which advocates for AI-driven digital transformation in the pensions sector ([8].

In theoretical literature, several theories underpin retirement decision-making. The Life Cycle Theory by (Modigliani & Brumberg, 1954) as quoted by [19] suggests that individuals plan retirement based on lifetime income expectations. The Push and Pull Theory by (Shultz et al., 1998) as quoted by [17] differentiates between external factors (e.g., workplace conditions) pushing individuals into retirement with internal motivations (e.g., financial security) pulling them into voluntary retirement. The Behavioral theories, such as Prospect Theory by

(Kahneman & Tversky, 1979) as quoted by [20], highlight how risk perception affects financial decisions related to retirement.

B. Application of ML in Retirement Prediction

With empirical literature, comparing ML and traditional statistical techniques in retirement prediction highlights ML's advantages in handling large datasets. Studies by [10]and [12] emphasize ML's ability to uncover hidden patterns, outperforming traditional parametric methods. ML models such as Logistic Regression, Random Forest, Support Vector Machines, have been tested for early retirement predictions, achieving accuracies between 70% and 97% ([21]; [22]). Zimbabwe's pension regulator, IPEC, advocates for AI integration in policy formulation, reinforcing the relevance of this study.

This study includes literature on ML applications in retirement predictions, pension analytics, and labor economics. Empirical studies utilizing ML algorithms such as Logistic Regression, Support Vector Machines, Random Forest, MLP and XGBoost are included in this study. Excluded are purely theoretical studies lacking ML implementation and research focusing solely on statistical techniques without computational modeling. Studies emphasizing early retirement due to attrition are included, given their relevance to retirement motives. Studies emphasizing delayed retirement have not been found and are the gap focus area that inspires the relevance of this study of retirement motives.

Literature highlights ML's superiority in early retirement prediction compared to traditional methods. While various ML models have been applied to early retirement motives, gaps remain in integrating diverse retirement motives into a single predictive framework. This study seeks to address these gaps by leveraging Neural Networks and Ensemble models to improve classification accuracy and enhance pension planning in Zimbabwe.

III. METHODOLOGY AND MATERIALS

A. Data Collection

The study utilized ML supervised classification models on a nationally representative dataset collected from members of a statutory pension scheme in Zimbabwe to predict retirement motives. The study compared five supervised learning classification models and selected the most robust for the prediction project. The dataset includes a wide range of variables such as age, gender, health status, years of service, financial literacy, and specific retirement goals. This data was used to train the models and evaluate their performance in predicting retirement motives.

The dataset used in this study was obtained from anonymized pension records of 4,411 employees in Zimbabwe, collected between 2020 and 2023 and stored in an Excel file. It included 30 features such as Age, TotalWorkingYears, JobSatisfaction, MaritalStatus, and FinancialReadiness. The target variable, RetirementDecision, is binary: early (1) or delayed (0) retirement. The dataset is moderately imbalanced, with 84% classified as delayed retirement motives and 16% as early retirement motives. All categorical features were encoded, and numerical features were scaled prior to model training. This dataset was well-suited for ML tasks aimed at predicting retirement motives due to its comprehensive coverage of demographic, financial, job centered and psychological factors.

B. Ethics and Data Consent

This study was conducted using anonymized administrative data obtained from a registered pension fund operating in Zimbabwe. All data were collected in accordance with Zimbabwean data protection laws and the fund's internal data governance policies. No personally identifiable information (PII) was used, and all sensitive variables (e.g., names, identification numbers) were removed or masked prior to analysis. The research team did not have access to direct identifiers at any stage of the analysis.

The study protocol was reviewed and approved by the Harare Institute of Technology (HIT) School of Information Science and Technology. The data provider granted explicit permission for the use of the dataset for academic research purposes, and all participants provided informed consent at the point of initial data collection for their data to be used in aggregate research analysis. This research complies with the ACM Code of Ethics and the principles of responsible data science, including fairness, accountability, and transparency.

C. Feature Engineering

Key features extracted from the raw data, to represent the factors influencing retirement decisions underwent data cleanup and engineering. With the acquired structured dataset from a statutory pension provider, data preprocessing, involving the handling of missing values, removal of duplicates, normalizing numerical features, and encoding categorical features signaled the setting in of exploratory data analysis (EDA). Data cleanup preceded the beginning of feature engineering through the extraction of a heat-map correlation matrix to identify key predictive variables.

After exploring the target variable, a massive class imbalance in the binary classification feature, was unearthed. This led to more explorations including the boxplots to expose outliers seen as another challenge in the dataset. Winsorization, to eliminate the outliers, Variance inflation Factor (VIF) and Principal Component Analysis (PCA) were invoked to reduce multi-collinearity. SMOTE was run to eliminate the class imbalance while StandardScaler standardization was carried out to normalize the features. The initial dataset had 30 features, but at the end of feature engineering only 21 remained including the target feature. Those rejected were either irrelevant or were redundant with a unique value or explained fully (to a large extent) by one or more other features (multi-collinearity). Having completed EDA ad with a balanced dataset on hand, model building commenced.

D. Data Split

The dataset, was split such that the training data was not used in either model validation or testing. The training-test split had a ratio of 70:30 with the majority data (70%) being utilized for training and the remaining (30%) shared equally for validation and testing. The validation-test split was 1:1 resulting in a final train-validation-test split ratio of 0.7:0.15:0.15. This ensured that 15% of the data was set aside for hyper-parameter tuning and the other 15% for testing the model’s ability to perform with unforeseen data. Extant literature agrees that a significant portion of the dataset, should be used for training. [11] and [3] used a 70:30 percent training-validation-test split. In previous studies the cross-validation (CV) method used was CV equals five (CV=5), but for this study CV equals three (CV = 3) was applied.

E. Hyper-parameter Optimization

Hyper-parameter tuning was performed to identify the parameters that would result in the optimal performance of each model. Table 1: Key Parameters, displays the optimal parameters that were tuned for each of the five models.

TABLE 1: Key Parameters

Model	**Baseline Accuracy**	**Key Parameters**
Logistic Regression	0.8550	{'C': 1.0, 'penalty': 'l2', 'solver': 'lbfgs',’max-iter’:’1000’}
Support Vector Machine	0.8535	{'C': ‘10’, 'kernel': ‘rbf’, 'gamma': ‘1’}
Random Forest	0.9849	{'n_estimators': 200, 'max_depth': ‘20’, 'criterion': 'gini'}
Multilayer Perceptron	0.9758	{'hidden_layer_sizes': (100,), 'activation': 'relu', 'alpha': 0.001,’max_iter:1000’}
XGBoost	0.9819	{'n_estimators': 300, 'max_depth': 7, 'learning_rate': 0.3}

F. Handling Class Imbalance

The target feature is a binary variable that relates to whether the employee intents to retire early or to delay retirement and continue working. In the dataset NO means motive is to delay retirement (0) or simply "Stay" while YES means motive is to retire early (1) meaning "Retire" or “Go”. The pie chart below in Figure 1: Target Feature

Distribution – RetirementDecision, shows that employees who intent to retire early were fewer (16.1%) than those who intent to delay retirement (83.9%) suggesting that the dataset was biased towards continuing work beyond retirement age. However, according to [23] algorithms will not give accurate results if the target variable has skewed class values. The distribution of the data reveals a massive imbalance of the target feature between the majority class "0" and the minority class "1". The distribution plotted in Figure 1: Target Feature Distribution depicts the class imbalance. This unbalanced distribution requires a solution to avoid misleading predictions.

According to [11] and [24], class imbalance arises when in a dataset the training data in one class is proportionally larger than the training data in the competing class. They posit that when a ML model is trained on imbalanced data, the trained model tends to capture the bias inherent in the classes. This causes the model performance metrics to be lower than if the classes were balanced before training.

To handle this imbalance, SMOTE was used in this study. [25] posits that SMOTE (Synthetic Minority Over-sampling Technique) is a technique used to handle imbalanced datasets in classification problems by generating synthetic samples for the minority class rather than simply duplicating existing ones. In this study's dataset, class (1) has significantly fewer instances than class (0) which is typical class imbalance. [25] further argue that models may become biased toward the majority class or misclassify minority class and techniques like SMOTE help by creating synthetic examples for the minority class, making the training set more balanced and improving model performance. After SMOTE the classes in this study, balanced at counts of 50% each as depicted in Figure 1: Target Feature Distribution - RetirementDecision (After SMOTE).

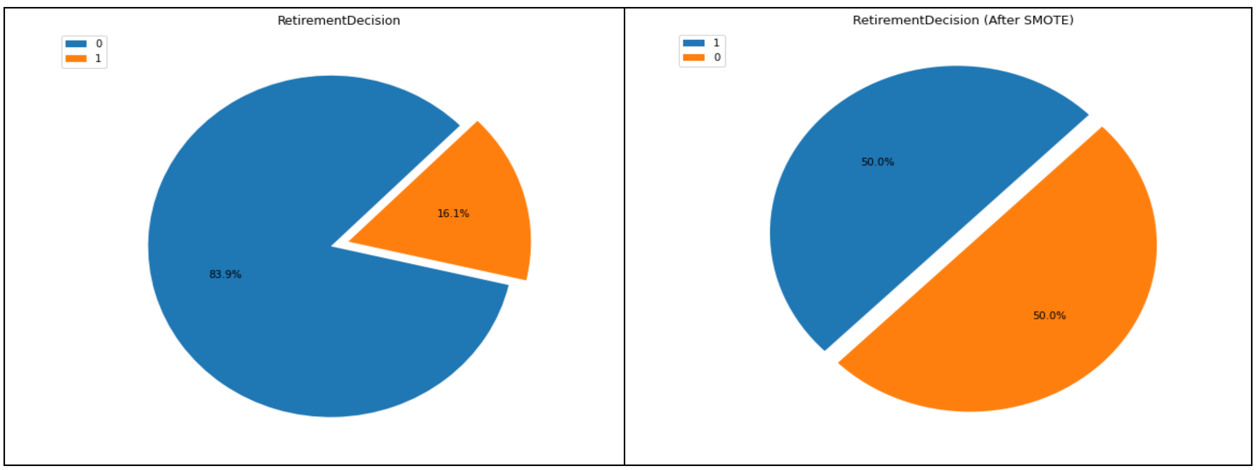

Figure 1: Target Feature Distribution

G. Model Training and Testing

Several ML models were selected for their suitability in handling classification problems, including Logistic Regression (LogReg), Random Forest (RF), Support Vector Machines (SVM), Multilayer Perception (MLP) and Extreme Gradient Boosting (XGBoost). The dataset, split into training and testing sets (70% for training and 30% for validation/testing) is the one utilized in model building. The trained modes tuned using GridSearchCV and cross-validation to optimize their performance, delivered acceptable results. The aim was to identify an optimal classifier that predicts the likelihood of employees staying in employment instead of retiring (binary retirement (Go) or non-retirement (Stay)).

Five ML classifiers were considered for predicting retirement motives including Logistic Regression (LogReg), Random Forest (RF), Support Vector Machines (SVM), Multilayer Perception (MLP) and Extreme Gradient Boosting (XGBoost). In the context of this study, outcome y refers to the predicted target feature under investigation whether the prospective retiree chose to continue working instead of leaving. This feature expressed in binary form can only take one of two possible values, "1" or "0". The value of "1" indicates that the participant has opted to "retire" while the value of "0" indicates that they have opted to "Stay" and continue working. The binary representation of the outcome simplifies the predicted target numeric value and facilitates classification computation given predictor values. Using the binary feature to represent the retirement motive researchers can easily compare and analyze the factors that contribute to or influence this outcome (Geron. A., 2019).

1) *Logistic Regression (LogReg) Classifier:* LogReg is a classification and regression model for binary prediction outcomes. According to [11] LogReg has shown high accuracy (over 80%) in several studies of early retirement prediction including the study carried out by [26] with a reported accuracy of 87%, [27] with a reported accuracy of 83%., [28] with a reported accuracy of 81%. In these studies, LogReg classifier outperformed other models including SVM, MLP, Decision Tree, KNN, Naïve Bayes and XGBoost. MLP only outperformed LogReg in the study by [22] of early retirement which included employee absenteeism and the accuracy rate was 78%. This made LogReg a good choice for joint early and delayed retirement binary classification in this study. According to [29] Logistic Regression is commonly used for classification, as it can output a value that corresponds to the probability (coefficient) of belonging to a given class. In this study LogReg achieved an accuracy of 85% which was below the 98% that Ensemble models Random Forest and XGBoost achieved.
2) *Support Vector Machine (SVM) Classifier:* With respect to SVM there is no evidence of high performance rating in previous studies with respect to the early retirement prediction domain. However, SVM is a very powerful and versatile ML Model [29], capable of performing linear or nonlinear classification, regression, and even outlier detection. The model SVM heavily favored the majority class (stay), leading to poor minority class performance. High precision for class 1 (73%) reveals that when the model says someone will retire, it is often correct. Very low recall (11%) for class 1, reveals that the model is missing most of the employees who actually retire. According to [29] this is a common issue in imbalanced classification tasks when baseline parameters are used. The base SVM model (C=10, kernel='rbf') would be the best to use in practice.
3) *Random Forest (RF):* In this study, RF exhibits outstanding performance, no false positives and only 20 false negatives. It is a best performer in this study revealing high precision, high recall, and excellent overall accuracy of 98%. According to [29], RF is an ensemble of Decision Trees, generally trained via the bagging method. [3] in an early retirement study of employee dissatisfaction motive, achieved an accuracy of over 90% reinforcing previous performances that ranked RF highly as an ensemble model that delivers top tier performance.
4) *Multilayer Perception (MLP) Classifier:* The MLP classifier was used for classification in this study. In previous studies [22] and [5] reported accuracy of over 90% with this neural network model. In this study, the baseline MLP model performed well on the majority class (class 0, non-retire motive) with very high recall (0.99), and high recall for class 1 (retire motive) at 0.90, including a modest F1-score (0.92) for retire motives showing that many true retire motives were not missed. The overall accuracy of 98% is inflated by the class imbalance (non-retire motives dominate the dataset). This outcome however underscores the importance of considering neural network applications, especially in imbalanced classification problems like employee retirement prediction.
5) *Extreme Gradient Boosting (XGBoost) Classifier:* This XGBoost model is outstanding in performance, striking an ideal balance between identifying retire motives and avoiding misclassification of non-retire motives. It is likely the best-performing model in this comparison and would be highly reliable for real-world retirement prediction tasks. In this study, it achieved a 98% accuracy. In previous studies by [30] in a study of employee attrition and [31] in a study of employee turnover, both recorded high accuracy of 93% and 97% respectively. This confirms the trend that ensemble models perform very well in the prediction of retirement motives.

IV. RESULTS AND DISCUSSION

The findings of this study underscore the potential of ML in enhancing the understanding of retirement motives among pension scheme members in Zimbabwe. The superior performance of the XGBoost, MLP and Random Forest models highlights the complexity and non-linearity of factors influencing retirement decisions. The importance of health status, job embeddedness and financial literacy in predicting retirement motives suggests that pension fund administrators should prioritize member education and health wellness programs. Moreover, personalized pension planning tools could be developed based on predictive modeling to assist members in making informed retirement decisions. The performance of the models was evaluated based on accuracy, precision score, recall score and F1-score the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve.

A. Descriptive Data Analysis

The dataset shape had 4411 rows and 30 columns. [29] suggests that it is generally important that the dataset is large enough (especially relative to the number of attributes), to avoid running the risk of introducing a significant sampling bias. In this case this retirement dataset is large enough for the 30 features. After preprocessing involving data cleanup and feature engineering, 20 features were retained for model building excluding the dependent variable.

The correlation between the predictor features and the target RetirementDecision feature area displayed in Figure 2: Correlation Heatmap. To determine the relationships among the features, Seaborn's heatmap function was employed for correlogram visualization. The results shown indicate that MaritalStatus had the most significant correlation with "retire" (class 1) motive followed by DecliningHealth and FamilyUpkeep. TotalWorkingYears had the most significant correlation with "Stay" (class 0) motive followed closely by YearsWithCurrentManager and YearsAtCompany. This finding is in line with intuitive notion that the longer the tenure or length of service the more likely an employee will stay longer.

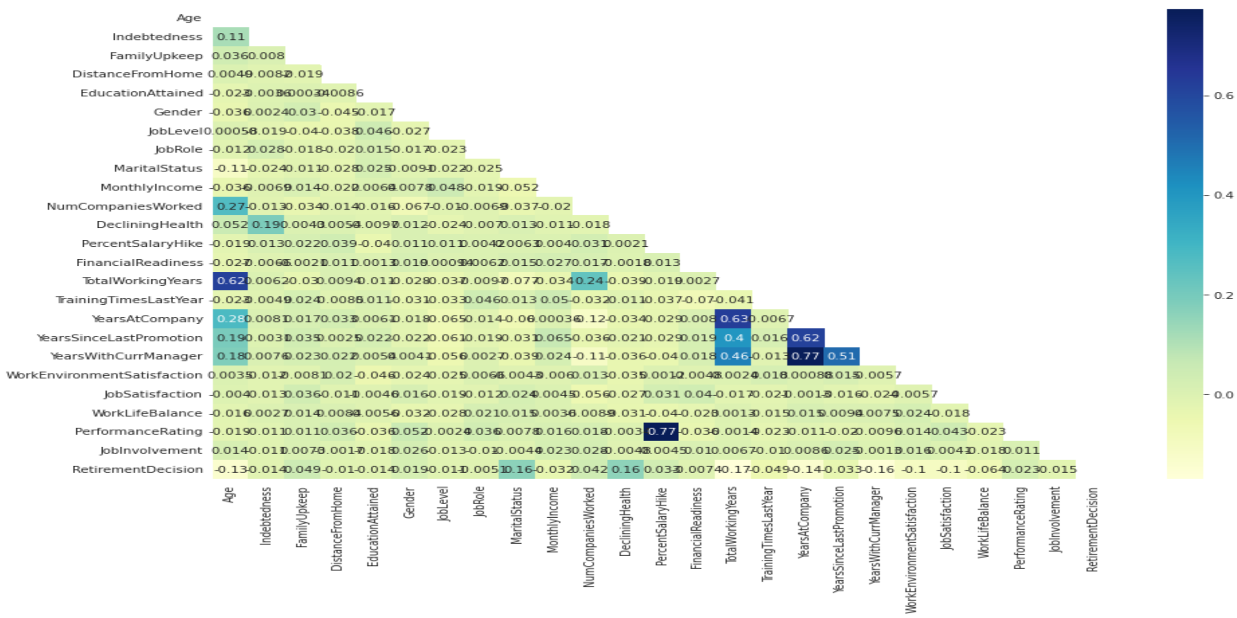

Figure 2: Correlation Heatmap

B. Evaluation Metrics on Up-Sampled Data

Table 2: Performance Comparison shows the evaluation metrics (accuracy, precision score, recall score and F1 score) from the various models' balanced data up-sampled using SMOTE. Based on these evaluation metrics, ensemble models XGBoost and Random Forest (accuracy 98%) show outstanding performance, striking an ideal balance between identifying retire motives and avoiding misclassification of non-retire motives. These best-performing models in this comparison would be highly reliable for real-world retirement prediction tasks. However, MLP is a viable alternative as its performance (98%) is overly acceptable, outperformed by a slight margin in all metrics for the retirement predictions task in the pensions sector. This dataset presents complex non-linear patterns and standalone models LogReg and SVM (85% accuracy) clearly struggle with the complex non-linear relationships among the features. Ensemble and neural network models however deliver excellent performances (98% accuracy).

TABLE 2: Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	Comment
LogReg	**0.8550**	**0.84**	**0.85**	**0.81**	Struggles with complex, non-linear relationships.
SVM	**0.8535**	**0.84**	**0.85**	**0.80**	Struggles with complex, non-linear relationships.
RF	**0.9849**	**0.99**	**0.98**	**0.98**	Learns complex non-linear patterns.
MLP	**0.9758**	**0.98**	**0.98**	**0.98**	Learns complex non-linear patterns.

Model	Accuracy	Precision	Recall	F1-Score	Comment
XGBoost	**0.9819**	**0.98**	**0.98**	**0.98**	Learns complex non-linear patterns.

C. Areas under Receive Operating Characteristic Curves

Figure 3: Area under the Curve (AUC), displays the areas under the curves (AUCs) of the five models in this study. The dotted diagonal line represents the 0.5 mark which is the point at which the AUC would be equivalent to a random guess or a fair coin toss. All AUCs represented by the five models are better than a random guess with Random Forest and XGBoost having a very high AUC of 0.99 in distinguishing between retirement uptake and non-uptake. The second highest AUC is shown by MLP with a value of 0.97 followed by Logistic Regression and SVM which both had 0.71.

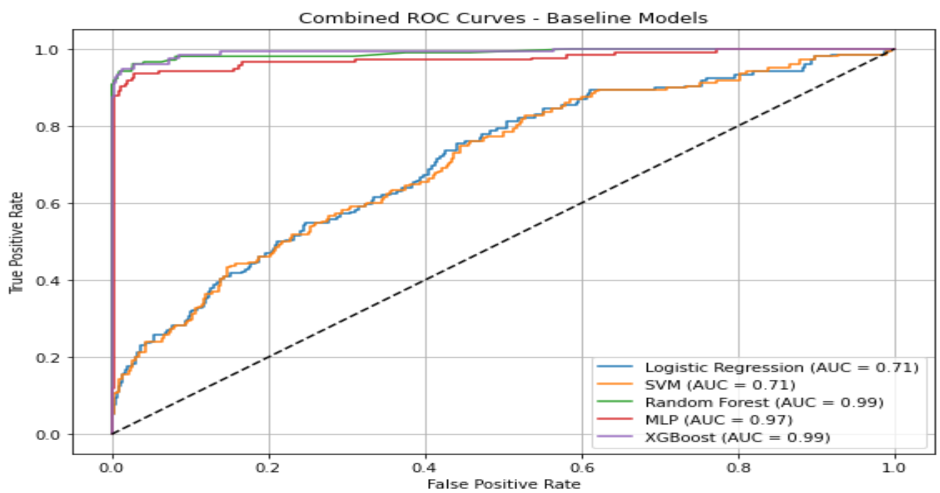

Figure 3: Area under the Curve (AUC)

These AUCs reveal that Random Forest and XGBoost are the most robust models in differentiating between the retirement uptake and non-uptake as these two have the highest area under the receiver operating characteristic curve. Logistic Regression and SVM with an AUC of 0.71 have a modest ability to distinguish between the binary classes (1 meaning retire and 0 meaning stay), better than random guess, but far from ideal. This matches the confusion matrix findings (poor recall). They are not reliable when identifying the positive class, very high false negatives. MLP with an AUC of 0.97, performs exceptionally well, just like Random Forest and XGBoost, very high accuracy in classification across thresholds, confirmed by high true positives and low false negatives in the confusion matrix.

This finding differs from studies by [27] and [28] who found out that standalone model Logistic Regression outperformed Random Forest in a study of the effect of economic and individual variables on retirement intentions. Based on the AUCs in Figure 3: Area under the Curve (AUC), the ensemble based and neural network models show better AUCs compared to standalone models. This finding is supported by [22] who found neural network model MLP performing better than standalone SVM and [3] who found ensemble based Random Forest recording a higher AUC than standalone models Logistic Regression and SVM.

The results of our study reveal that XGBoost and Random Forest with an AUC of 0.99 are excellent classifiers achieving almost perfect distinction between classes. They are extremely reliable models for classification with ROC curve which matches confusion matrix showing a very few false positives and very few false negatives. These two are top-tier performance models, reliable across all thresholds, excelling at both detecting positives and minimizing false alarms.

D. Confusion Matrices

The performances of the models in the retirement motive prediction study, were measured using the confusion matrix evaluation tool. Table 3: Confusion Matrices Summary displays confusion matrix results relating to the models in this study. The results show LogReg with low false positives (only 10). Very high false negatives (182), meaning it misses most positive cases. This model is biased towards the negative class, likely due to class imbalance. It shows poor recall for the positive (Retire) class. SVM shows slightly fewer false positives than LogReg. Even worse recall (only 22 true positives), with the highest false negatives (186). Similar to LogReg, this model struggles to detect positive (Retire) cases. Very low sensitivity/recall. Random Forest shows an outstanding performance, no false positives and only 20 false negatives. Probably the best performer in this group exhibiting high precision, high recall, and excellent overall accuracy. MLP matches Random Forest in detecting positive cases with true positives equal to 188. Slightly more false positives than Random Forest. Very good performance, comparable to Random Forest with a small drop in precision. XGBoost ranked as the best model shows low recall with 20 false negatives, high true positives of 190, and low false positives equal to 6. This model exhibits best balance between sensitivity and specificity and is likely the overall best model for both classes. XGBoost and Random Forest clearly outperform the rest, especially in handling the positive class. LogReg and SVM perform poorly, likely due to class imbalance, consideration should be made for using resampling or class weighting for improvement. MLP is a strong contender but slightly less precise than Random Forest and XGBoost.

TABLE 3: Confusion Matrices Summary

Model	**TP**	**FP**	**FN**	**TN**	**Comment**
LogReg	26	10	182	1106	Poor recall, bad for minority class
SVM	22	8	186	1108	Even worse recall, bad for minority class
Random Forest	188	0	20	1116	High precision and recall balanced classes
MLP	188	12	20	1104	High recall, slightly lower precision
XGBoost	190	6	18	1110	Best class balance; top performer overall

According to [18], the confusion matrixes are constructed considering a threshold of 0.5. This means that if the predicted probability of the event in the test set was greater than 0.5, it classified as “Retire”, otherwise “Stay”. The numbers along the main diagonal represent the correct decisions made in the test set. From the Table 3: Confusion Matrices Summary, it can be obtained that the decisions made correctly were 1132, 1130, 1304, 1292 and 1300 for logistic regression, SVM, random forest, MLP and XGBoost, respectively. Although there is not much difference between these values, it is observed that Random Forest has the best performance followed closely by XGBoost, when predicting true events of early and delayed retirement.

This Random Forest and XGBoost models show outstanding performance, striking an ideal balance between identifying retire motives and avoiding misclassification of non-retire motives. They are the best-performing models in this comparison and would be highly reliable for real-world retirement prediction tasks. However, MLP is a viable alternative as its performance is overly acceptable for the retirement predictions in the Pensions Sector.

E. Feature Importance

The top features provide a coherent narrative: job embeddedness, stability, and career growth are central to the “Stay” motive. This insight is critical for proactive retirement and HR policy and strategy. Figure 4: Ranked Feature Importance displays the top 15 features that influence retirement decisions in this study.

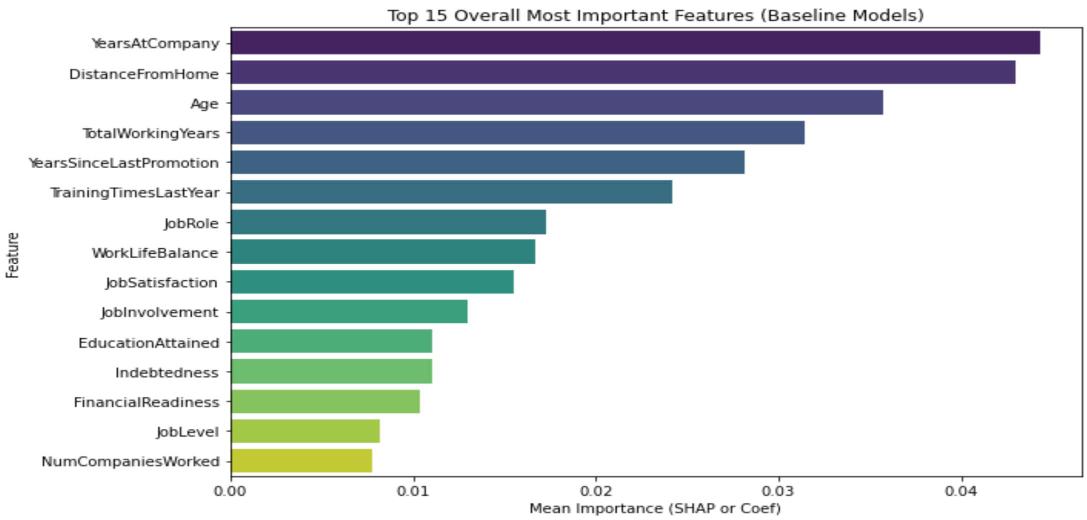

Figure 4: Ranked Feature Importance

Advanced models such as MLP, Random Forest, and XGBoost excelled due to their ability to learn complex non-linear decision boundaries rather than relying on class-wise linear coefficients. These models assess feature influence holistically. This is evident in the overall feature importance rankings derived from SHAP (SHapley Additive exPlanations) values in this study.

The top-ranked features provide a coherent narrative supporting the "Stay" motive, including factors like job embeddedness, stability, and career growth. These insights have significant implications for developing proactive HR and retirement policies. The top 15 influential features, are displayed in TABLE 4: Top 15 Overall Important Features

TABLE 4: Top 15 Overall Important Features

Rank	**Feature**	**Importance**
1	YearsAtCompany	0.044348
2	DistanceFromHome	0.042976
3	Age	0.035709
4	TotalWorkingYears	0.031418
5	YearsSinceLastPromotion	0.028184
6	TrainingTimesLastYear	0.024154
7	JobRole	0.017321
8	WorkLifeBalance	0.016700
9	JobSatisfaction	0.015539
10	JobInvolvement	0.013022
11	EducationAttained	0.011064
12	Indebtedness	0.011040

Rank	Feature	Importance
13	FinancialReadiness	0.010373
14	JobLevel	0.008181
15	NumCompaniesWorked	0.007757

Figure 5: Top Feature Overlap, illustrates the feature overlap and consistency across models based on the top 15 features selected by each model. The significant overlap among these feature sets demonstrates a converging understanding across models regarding the most influential predictors of the retirement decision. This convergence increases confidence in the reliability of the feature selection process.

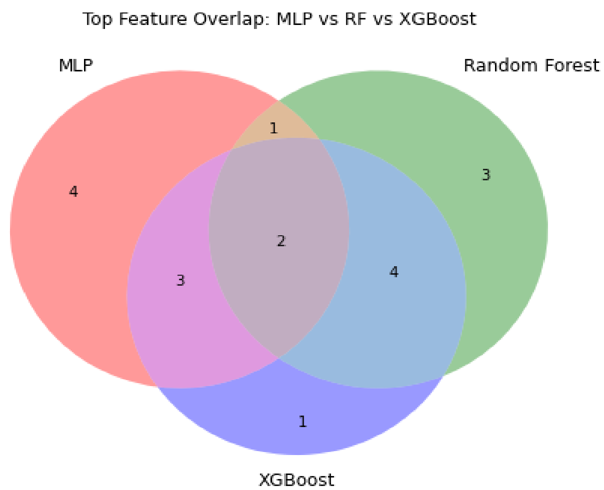

Figure 5: Top Feature Overlap

F. Comparison with previous study results

According to [11], Logistic Regression (LogReg) has demonstrated high accuracy (above 80%) in several early retirement prediction studies. Notably, [26] reported an accuracy of 87%, [27] reported 83%, and [28] achieved 81%. In these cases, Logistic Regression outperformed other models such as SVM, MLP, Decision Trees, K-Nearest Neighbors (KNN), Naïve Bayes, and even XGBoost. However, MLP outperformed Logistic Regression only in the study by [22], which focused on early retirement prediction in the context of employee absenteeism, with an accuracy of 78%.

These findings positioned Logistic Regression as a reliable model for both early and delayed retirement prediction tasks. Additionally, Random Forest was rated the best classifier in the study by [3], with accuracy exceeding 80%, while Decision Tree was identified as the top performer in an employee attrition study by [21], achieving an impressive 97% accuracy. On the other hand, SVM has not consistently shown strong performance in the retirement prediction domain based on existing literature. The MLP model, a neural network capable of capturing non-linear patterns, was also recognized by [22] as the best-performing model in a study involving long-term absentee employees, achieving 78% accuracy. Similarly, XGBoost, according to [32], is widely acknowledged as a powerful ensemble learning algorithm, though exact accuracy metrics were not reported in their findings.

While previous studies highlighted the strong performance of Logistic Regression, its performance in the current study was relatively modest, achieving an accuracy of 85%, similar to SVM, which also recorded 85%. These results contrast with earlier research where Logistic Regression often outperformed other models. In contrast, the Random Forest, MLP, and XGBoost models significantly outperformed both Logistic Regression and SVM in this study, each achieving an accuracy of approximately 98%. These models benefit from their capacity to capture complex, non-linear relationships and adapt to intricate patterns in the data. The disparity between the performance of Logistic

Regression in this study versus prior work could be attributed to methodological differences. Notably, this study utilized SHAP (SHapley Additive exPlanations) for feature importance analysis rather than relying solely on probability coefficients or raw model weights. The use of mean SHAP values may have introduced a different interpretative lens, possibly contributing to the reduced performance of linear models like Logistic Regression and SVM.

Overall, Logistic Regression, though historically strong, underperformed in this study compared to non-linear and ensemble methods. XGBoost, MLP, and Random Forest outperformed other models, showing strong alignment with recent trends favouring models capable of handling complex feature interactions. The use of SHAP-based feature importance may have contributed to the observed differences in performance, particularly for models that assume linear separability. This study highlights the evolving strength of advanced models like XGBoost and MLP in domains traditionally dominated by interpretable models such as Logistic Regression.

G. Critical Analysis of Findings

XGBoost, Random Forest, and MLP outperformed other models by effectively balancing retirement motive prediction accuracy with minimal misclassification. XGBoost led overall, followed closely by MLP and Random Forest, making all three strong candidates for deployment in the pensions sector. Unlike linear models such as Logistic Regression and SVM, these top models capture complex non-linear patterns without relying on class-specific coefficients. While this limits interpretability (e.g., no direct indication of whether a feature increases or decreases a class outcome), they still offered coherent insights especially with features like job embeddedness, career growth, and stability strongly linked to the “Stay” motive.

All three models achieved ROC-AUC scores above 0.97, showing excellent class separation. XGBoost delivered the best results with 98% accuracy, F1-score, lowest false negatives (18), and highest true positives (190). MLP and Random Forest also performed well, though slightly less balanced. In contrast, Logistic Regression and SVM lagged likely due to class imbalance though they offer greater transparency and could benefit from techniques like resampling or class weighting. Given the trade-off between performance and interpretability, XGBoost is recommended for deployment, with Random Forest and MLP as viable alternatives depending on context. MLP also showed strong generalization, while SVM and Logistic Regression require tuning before use.

This study confirms that ML can effectively predict retirement motives in Zimbabwe, with Random Forest and XGBoost surpassing traditional models. Key predictors like Age, Health Status, Work-Life Balance, and Years at Company were consistently significant. ML's strength in handling non-linear, imbalanced data positions it as a superior tool for strategic pension planning. Stakeholders like IPEC, NSSA, and pension providers should adopt ML based forecasting for improved policy accuracy and responsiveness. Tailored pension products can be developed using insights from model outputs (e.g., SHAP values). Addressing data imbalance with techniques like SMOTE will further enhance fairness and performance. There is a strong case for expanding AI-driven retirement research across sectors, integrating NLP for unstructured data, and aligning with Vision 2030 by embedding AI tools into the national pension ecosystem to modernize services and support retiree welfare.

V. CONCLUSION

This research demonstrates the applicability and benefits of using ML techniques to predict retirement motives in the Zimbabwean pensions sector. By leveraging advanced analytics, pension fund administrators can gain deeper insights into member needs and preferences, ultimately leading to more effective pension scheme management and improved member satisfaction. Future research could explore the integration of additional data sources, such as economic indicators and social media data, to further enhance predictive performance. ML's ability to process and interpret large, complex datasets positions it as a transformative tool for stakeholders in Zimbabwe's pension sector, including policy makers, regulators, pension funds, and employees themselves. Using supervised learning models such as Logistic Regression, Support Vector Machine (SVM), Random Forest, Multi-Layer Perceptron (MLP), and Extreme Gradient Boosting (XGBoost), the study assessed prediction accuracy, model interpretability, and feature influence to determine what drives retirement decisions. These insights are expected to improve fund planning and regulation, support targeted retirement products, and ultimately align pension strategies with individual realities and national economic objectives.

REFERENCES

[1] D. A. M. Figueira, M. do C. L. Haddad, R. Gvozd, and P. de S. C. Pissinati, "Retirement decision-making influenced by family and work relationships," *Rev. Bras. Geriatr. e Gerontol.*, vol. 20, no. 2, pp. 206–213, 2017, doi: 10.1590/1981-22562017020.160127.

[2] T. A. Beehr and M. M. Bennett, "Working after retirement: Features of bridge employment and research directions," 2015. doi: 10.1093/workar/wau007.

[3] A. Patel, N. Pardeshi, S. Patil, S. Sutar, R. Sadafule, and S. Bhat, "Employee attrition predictive model using machine learning," *Int. Res. J. Eng. Technol.*, vol. 7, no. 5, pp. 3855–3859, 2020.

[4] D. Kooij, A. de Lange, P. Jansen, and J. Dikkers, "Older workers' motivation to continue to work: Five meanings of age: A conceptual review," 2008. doi: 10.1108/02683940810869015.

[5] Y. Sun and H. Jung, "Machine Learning (ML) Modeling, IoT, and Optimizing Organizational Operations through Integrated Strategies: The Role of Technology and Human Resource Management," *Sustain.*, vol. 16, no. 16, 2024, doi: 10.3390/su16166751.

[6] G. Irlam, "Machine learning for retirement planning," *J. Retire.*, vol. 8, no. 1, pp. 32–39, 2020, doi: 10.3905/JOR.2020.1.067.

[7] M. N. News, Local, "Revolutionising short-term insurance in Zimbabwe : The AI advantage," 2024.

[8] B. K. M. A. T. Mazikana, P. Hove, "ZimVision2030," 2024.

[9] J. Oakman and Y. Wells, "Retirement intentions: What is the role of push factors in predicting retirement intentions?," 2013. doi: 10.1017/S0144686X12000281.

[10] M. G. Garibay, A. Srakar, T. Bartolj, and J. Sambt, "Does Machine Learning Offer Added Value vis- à- vis Traditional Statistics? An Exploratory Study on Retirement Decisions Using Data from the Survey of Health, Ageing, and Retirement in Europe (SHARE)," *Mathematics*, vol. 10, no. 1, 2022, doi: 10.3390/math10010152.

[11] N. Kemboi Yego, J. Kasozi, and J. Nkurunzinza, "Optimizing Pension Participation in Kenya through Predictive Modeling: A Comparative Analysis of Tree-Based Machine Learning Algorithms and Logistic Regression Classifier," *Risks*, vol. 11, no. 4, 2023, doi: 10.3390/risks11040077.

[12] Kantara F. R., "Predicting Early Retirement Intentions of Mature Workers using Machine Learning and Deep Learning Algorithms," *Tilbg. Univ.*, 2021.

[13] T. M. Amabile, "Amabile Retiment -Aging.pdf," 2019, *Harvard Business School*.

[14] D. B. I. P. V. Baherec, "Barbic Machine Learning.pdf," 2016, *University of Zagreb*.

[15] B. Seligman, S. Tuljapurkar, and D. Rehkopf, "Machine learning approaches to the social determinants of health in the health and retirement study," *SSM - Popul. Heal.*, vol. 4, no. July 2016, pp. 95–99, 2018, doi: 10.1016/j.ssmph.2017.11.008.

[16] S. Espinoza, A. Malhotra, and D. Zhang, "Predicting Chilean Retirement Plan Decisions with Deep Neural Networks," *Stanford G.S.B.*, pp. 2–6, 2022, [Online]. Available: http://cs230.stanford.edu/projects_fall_2020/reports/55775397.pdf

[17] C. M. A. Alessia Negrini l, Chiara Panari, Silvia Simbula, "Journal of Work and Organizational Psychology," *J. Work Organ. Psychol.*, vol. 30, no. 2, pp. 51–58, 2013.

[18] J. de Jesus Rocha Salazar, "The Effect of Economic and Individual Variables on Retirement Decisions," *Dissertation*, no. May, 2019.

[19] Z. Bodie, J. Treussard, and P. Willen, "The Theory of Life-Cycle Saving and Investing," *SSRN Electron. J.*, no. February, 2011, doi: 10.2139/ssrn.1002388.

[20] H. Sonnenschein, " Econometrica ," *Encycl. Stat. Sci.*, vol. 0100, no. 3469, pp. 263–291, 2005, doi: 10.1002/0471667196.ess0533.

[21] Y.-D. Yun, S.-H. Lee, H.-S. Ji, and H.-S. Lim, "Development of Retirement Prediction Model based on Work Life Profile Using Machine Learning Method," *J. Korean Assoc. Comput. Educ.*, vol. 20, no. 1, pp. 87–97, 2017.

[22] E. Lima, T. Vieira, and E. de B. Costa, "7 - Evaluating deep models for absenteeism prediction of public security agents.pdf," 2020.

[23] Josh Hugh Learning, "Python Machime Learning," p. 178, 2019.

[24] Provost & Fawcett, "Data science-what you need to know about analytic-thinking and decision-making," *J. Chem.*

Inf. Model., vol. 53, no. 9, pp. 1689–1699, 2013.
[25] X. Guo, Y. Yin, C. Dong, G. Yang, and G. Zhou, "On the class imbalance problem," *Proc. - 4th Int. Conf. Nat. Comput. ICNC 2008*, vol. 4, pp. 192–201, 2008, doi: 10.1109/ICNC.2008.871.
[26] J. Frierson and S. Dong, "Who's Next - Evaluating Attrition with Machine Learning Algorithms and Survival Analysis.," 2018.
[27] J. de J. R. Salazar and M. del C. B. Penas, "Scoring and Prediction of Early Retirement using Machine Learning Techniques: Application to Pension Plans," pp. 1–18, 2019.
[28] S. Najafi-Zangeneh, N. Shams-Gharneh, A. Arjomandi-Nezhad, and S. H. Zolfani, "4 - An Improved Machine Learning-Based Employees Attrition Prediction Framework with Emphasis on Feature Selection.pdf," 2021.
[29] A. Géron, *Hands-on Machine Learning whith Scikit-Learing, Keras and Tensorfow*. 2019.
[30] H. Alqahtani, H. Almagrabi, and A. Alharbi, "Employee Attrition Prediction using Machine Learning Models: A Review Paper," *Int. J. Artif. Intell. Appl.*, vol. 15, no. 2, pp. 23–49, 2024, doi: 10.5121/ijaia.2024.15202.
[31] Z. Yin, B. Hu, and S. Chen, "Predicting Employee Turnover in the Financial Company : A Comparative Study of CatBoost and XGBoost Models Predicting Employee Turnover in the Financial Company : A Comparative Study of CatBoost and XGBoost Models," 2024, doi: 10.20944/preprints202410.0072.v1.
[32] X. Chen, S. Lin, Y. Zheng, L. He, and Y. Fang, "Long-term trajectories of depressive symptoms and machine learning techniques for fall prediction in older adults: Evidence from the China Health and Retirement Longitudinal Study (CHARLS)," *Arch. Gerontol. Geriatr.*, vol. 111, no. August, pp. 1–8, 2023, doi: 10.1016/j.archger.2023.105012.

