Using the Ensemble Method to Predict Employee Attrition at the Multination Consulting Corporation

Molefe J Maleka[1]

[1]Office of DVC: Teaching and Learning, eMalahleni Campus, Tshwane University of Technology, eMalahleni, South Africa, Orcid: 0000-0003-1929-8532

Keywords

Attrition
Consulting organisation
Ensemble method
Remuneration
Work attitude

Abstract

The paper aimed to identify essential features that predict employee attrition at the multination consulting corporation. Losing talented employees is costly to organisations, as they incur recruitment and development costs and lose their competitive edge. Mobley, Griffeth, Hand and Meglino's (1979) model was used to give theoretical grounding to the study. The study was quantitative, and the positivism paradigm influenced it. It used secondary data and bivariate analysis, including correlation and bagging (i.e., bagging classifier, random forest, decision tree, weighted bagging classifier, and decision tree estimator), to analyse data and decide which model best predicts attrition.

The data showed the top significant predictors of employee attrition: monthly income, overtime, daily rate, age, hourly rate, distance from home, number of companies worked and work-life balance. This study contributes by using the emerging ensemble method in predicting the features that predict employee attrition. Since 4 of the top 8 predictors are remuneration-related, managers should focus on revising and benchmarking best practices in the remuneration policy. Management should also focus on developing strategies to retain employees between 18 and 28 years.

[1]Malekam@tut.ac.za

1. Introduction

1.1 Background

Investigating features that predict employee attrition is a topic of interest to scholars (Adriano & Callaghan, 2020) and human resource management (HRM) practitioners (Hancock, Allen, Bosco, McDaniel & Pierce, 2013). The former wrote academic papers and supervised students to develop models or frameworks that are published either in journal articles or conference proceedings. HRM practitioners know that if the organisation does not establish retention strategies to mitigate and reduce employee attrition, they spend a lot of money recruiting and selecting employees to replace those who have resigned (Bussin, 2018). HRM practitioners know that their organisations have lost their competitive edge in their industries (Liu & Raghuram, 2021). Gross (2022) mentioned that the cost to replace an employee who has churned, according to the Society of HRM, is between 30% to 40% for employees at the entry level, 125% at middle management and 200% at senior management.

Employee attrition is a complex and current topic because it is driven by many predictors (Tshwane, Maleka & Tladi, 2023), which differ per economic sector. This study focuses on the consulting sector because if they lose an experienced consultant, they can take clients with them. This can negatively impact the consulting company's revenue (Pratt, Boudhane & Cakula, 2021). This shows that it is in the consulting organisation's best interest to understand the factors leading to employee attrition and develop strategies to mitigate them (Dutta & Bandyopadhyay, 2020). A study in Indonesia in a consulting organisation showed that factors like poor reward and compensation, lack of training opportunities, and job satisfaction were the main factors contrition to employee attrition (Fakhri, Nurnida, Winarno & Suryana, 2020). The commonality between Pratt et al. and Duta and Bandyopadhyay used machine learning to identify which predicts attritions and such factors are external or what Herzberg called hygiene factors (Maleka, 2012) and what Mobley, Griffeth, Hand and Meglino (1979) called personal and work-related factors. Some of those predictors are external (i.e. trade unions, employment rates), work-related (i.e. remuneration, job performance and satisfaction) and personal (i.e. age, gender, marital status).

Whereas some of the attrition studies (for example Pratt et al., 2021) discussed above has used machine learning methods to predict employees, a survey in the articles and conferences published in the South African Journal of HRM (SAJHRM) and International Business Conference such approach is limited. South African scholars have used structural equation modelling (SEM) and Statistical Packages for Social Science (SPSS) in developing attrition models. The contribution of this study is that it also included environment satisfaction, which current South African literature did not focus on (Dhanpat & de Braine, 2022; Schaap & Olckers, 2020). The rationale of including it other that South African studies, is that internationally it had found employees who worked in the buildings that are not safe and bad working condition are most likely resign (Shin, 2016). As much as this research contributed to the turnover body of knowledge, it was based on a single data collection method using a few validated scales. Conducting employee attrition using SEM and SPSS is imitating because it relies on a few features to predict employee attrition (Oswald, Behrend, Putka & Sinar, 2019) to develop a single employee attrition model. And yet, in the literature (Cotton & Tuttle, 1986), it has been found that there are 26 predictors of employee attrition. Others have found 35 predictors of employees in consulting and other industries (Fallucchi, Coladangelo, Giuliano & William De Luca, 2020). It is envisaged that this study uses an insights ensemble method that compares different models to determine which is performing better in predicting employee attrition.

1.2. Problem statement

The information in the introduction shows that employee attrition is predicted by several factors (Dutta & Bandyopadhyay, 2020). In the multination consulting organisation where this study was conducted, managers were unaware of the predictors of employee attrition, and they were aware that it is costly to acquire and develop talent because they were operating in the consulting industry. In the consulting industry, attrition is at 38%, driven mainly by salary and benefits (Enrich, 2023). Hence, consulting managers are keenly interested in attrition and always use data to develop actionable retention strategies to mitigate attrition. However, such research is not conclusive. Hence, this study is salient to bring known insights into the attrition body of knowledge.

1.3. Research objectives

The objectives of the study are as follows:

- To determine the relationship between the attrition predictors;
- To compare existing machine learning models that predicts employee attrition; and
- To select appropriate machine learning models that predicts employee attrition.

As can be observed from Table 1 that turnover can be voluntary and involuntary. Mello (2015) opined that the former is when the employees leave an organisation at their initiative, and the latter is when an employee the organisation at the organisation's request. The other observation is that many predictors predict the organisation. The paper aligns itself with Fallucchi et al. (2020) definition as it covers predictors measured in this study. As previously discussed in the introduction, it is that employee attrition research should include many predictors. The theoretical framework, including personal, organisational and external predictors, is discussed in the next section.

2. Literature review

Employees' attrition definitions are shown in Table 1.

Table 1: Employees' attrition definitions

Definition	Author
Employee turnover is the leakage of intellectual capital from the workplace.	Punnoose and Ajit (2016)
Employee attrition is about an employee withdrawing from the organisation/s due to predictors like job satisfaction, personal employee characteristics and perceived alternative employment opportunities.	Michaels and Spector (1982)
It is defined as employees terminating their labour from the organisation willingly, without being pressured employer and due to lack of commitment and dissatisfaction.	Mathieu, Fabi, Lacoursière and Raymond (2016)
It is losing a permanently employed worker who has to be replaced by the organisation.	Grobler, Wärnich, Carell, Elbert & Hatlield (2011)
Employee attrition is about an organisation laying off employees due to financial constraints, and it can also include the HRM division implementing strategies that would ensure poor-performing employees are leaving. The high-performing employees stay within the organisation. Such strategies can include hiring access labour, training programmes and creating a core of talented and committed employees.	Mowday (1984)
Attrition can happen when an employees retire and resign from an organisation. Some predictor that leads to employees leaving is the distance from home, number of previous employers, hourly rate, business travel, stock options level, education, overtime, job roles and involvement, gender, age, percent salary hike, monthly income, overall and relations satisfaction, work-life balance and years in current role.	Fallucchi et al. (2020)

Source: Author's own construction

Attrition theoretical framework

Mobley, Griffeth, Hand and Meglino (1979) theoretical grounding. Even though this model has been replicated and widely used in the HRM (Behravesha, Tanovaa & d Abubakar, 2020) and industrial and organisational psychology literature (Michaels. & Spector, 1982), it only focuses on voluntary attrition. Similarly, this paper focuses on voluntary attrition, and it was empirical and unlike Mobley et al. (1979) attrition turnover, which was conceptual. Mobley et al. (1979) attrition framework is an update on Porter and Steers's (1973) literature analysis.

Like Porter and Steers analysis, Mobley et al. framework states that overall job satisfaction negatively affects employees. Recently, South African studies found the same correlations (Tshwane et al., 2023; Winarno, Prasetio, Luturlean & Wardhani, 2022). Job satisfaction is a multifarious construct comprising physiological, environmental, and psychological circumstances (Spector, 1997; Tirta & Enrika, 2020). Mobley et al. argued that even job dissatisfaction accounts for 16% of employee attrition. Factors that reduce job satisfaction include offering employee's autonomy, supervisory support, employee involvement in business unit decisions (Skosana, Maleka & Ngonyama-Ndou, 2021) and union membership. Unions are effective for negotiating concussive working conditions and higher salary increases (Maleka, 2018),

In this study, one of the work-related factors investigated is work-life balance, and this variable was not included in e Mobley et al. model. "Work-life balance has been defined as the degree to which there is equitable satisfaction and engagement relating to an employee's work and family roles" (Adriano & Callaghan, 2020, p.3). Work-life balance interests managers in the consulting industry due to the rising number of Generation Z who require flexibility. In addition, the work-life balance increment is driven by the increment of females in the workplaces, "the rising number of couples both working outside the home or the transformation of family structures as well as population ageing, technological advances, birth rate decline and the need to improve human capital management" (Benito-Osorio, Muñoz-Aguado & Villar, 2014, p. 214).

Mobley et al. framework state that job-related factors (i.e. organisational climate, practices and rewards) propel an employee to attrite or not. It has been established that in an organisation where the manager created a positive organisational climate, instead of churning, employees were happy and assisted the manager in reaching strategic objectives (Grobler & Jansen van Rensburg, 2019). In such organisations, employees have positive attitudes and affection towards organisational strategy (Castro & Martins, 2010). In the workplace, where employees were given training opportunities and managers were fair in appraising them, such employees were motivated and did not leave (Memon, Salleh, Mirza, Cheah, Ting, & Tariq, 2021). The literature is consistent that employees who are paid and are offered excellent benefits are least likely to resign, and the opposite is also true (Barkhuizen, Lesenyeho & Schutte, 2020).

Both Porter and Steers analysis, Mobley et al. frameworks showed individual factors like age, tenure, and gender play a role in employee attrition. In addition, it has been established that individual factor employees with the highest education levels are most likely to resign. The same study showed that employees who were attention seekers and were not receiving it resigned (Hughes, O'Brien, Reeder & Purl, 2020).

Mobley et al. framework posits that expectation also plays a critical roles in determining employees' turnover. This narrative emanated from studies that were Vroom's Expectancy Theory (Akgunduz, Gök & Alkan, 2020). This theory has three dimensions, viz: valence, instrumentality and expectancy. Valence is the desirability of potential outcomes (such as high pay) associated with a particular job. Instrumentality is the likelihood that the desired outcomes will be obtained from the job. Expectancy is about expecting somebody looking for an appointment to find it (Nouri & Parker, 2019). Vroom claimed that employees with realistic expectations and congruency between expectations and actual employees are least likely to leave. This can be in payment and relationship with the supervisor, manager and co-workers. In addition, Mobley et al. opined that giving employees practical information during orientation played a role in diminishing their propensity to attrite. Lastly, Mobley et al. talk about the labour market economics factors. One of them is unemployment, which is that employees are least likely to churn if there are vacancies in the labour market. Other labour economics factors propelling attrition of employees are job opportunities, skills demand, cost of living, quality of life, transport and health facilities (Al Mamun & Hasan, 2017).

It needs to be stressed that various attrition frameworks have been developed before and after the Mobley et al. seminal work (Hom, Lee, Shaw & Haushnecht, 2017). The trend in the literature is that other scholars opted to select affective and attitudinal predictors (i.e. employees' engagement, job satisfaction, organisational commitment) and conduct SEM to either confirm or falsify their relationship with attrition. Such research is mainly correlational and uses validated scales of attrition as the outcome variables (Dhanpat & de Braine, 2022). This research did not include personal predictors and labour economics in their analysis. The other trend, rare in HRM literature, is found mainly in computer science literature; it is the ensemble technique (Falluchi et al., 2020). The models they developed are either regression or binary based. This kind of changing focus on confirmatory research is advocated by Oswald et al. (2019), and they claimed that it should be adopted by industrial and organisational psychology and HRM scholars. They say that data is becoming big, many predictors should be used, and data can be assessed from different sources. In addition, they opined that as much as using SEM is useful, it requires few variables and gives a reductionist picture for HRM practitioners.

3. Methodology

3.1. Approach and paradigm

This study was quantitative, and the positivism paradigm influences it. Within the positivism paradigm, knowledge is created by conducting statistical techniques (Saunders, Lewis & Thornhill, 2019). The author of the paper followed the ensemble paradigm, which about used different models or what is called learners, and the researcher trains them and selects the one performing better to solve organisational problems (Zounemat-Kermani, Batelaan, Fadaee & Hinkelmann, 2021). Unlike in the SEM paradigm, where hypotheses are tested to confirm the theory, learner features/predictors are ranked according to their importance in the ensemble paradigm.

3.2. Data collection and sampling distribution

This study used secondary data from a global consulting company with around 3000 employees. The data collection instrument used has these variables shown below: Employee Number - Employee Identifier:

- Attrition - Did the employee attrite, 0 did not attrite and 1 attrite
- Age - Age of the employee
- Business Travel - Travel commitments for the job
- Daily Rate - Data description not available
- Department - Employee Department
- Distance From Home - Distance from work to home (in km)
- Education - 1-Below College, 2-College, 3-Bachelor, 4-Master,5-Doctor
- Education Field - Field of Education
- Employee Count - Employee Count in a row
- Environment Satisfaction - 1-Low, 2-Medium, 3-High, 4-Very High
- Gender - Employee's gender
- Hourly Rate - Data description not available**
- Job Involvement - 1-Low, 2-Medium, 3-High, 4-Very High
- Job Level - Level of job (1 to 5)
- Job Role - Job Roles
- Job Satisfaction - 1-Low, 2-Medium, 3-High, 4-Very High
- Marital Status - Marital Status
- Monthly Income - Monthly Salary
- Monthly Rate - Data description not available*
- Number Companies Worked - Number of companies worked at
- Over18 - Over 18 years of age
- Over Time - Overtime
- Percent Salary Hike - The percentage increase in salary last year
- Performance Rating - 1-Low, 2-Good, 3-Excellent, 4-Outstanding
- Relationship Satisfaction - 1-Low, 2-Medium, 3-High, 4-Very High
- Standard Hours - Standard Hours
- Stock Option Level - Stock Option Level
- Total Working Years - Total years worked
- Training Times Last Year - Number of training attended last year
- Work Life Balance - 1-Low, 2-Good, 3-Excellent, 4-Outstanding
- Years At Company - Years at Company

- Years In Current Role - Years in the current role
- Years Since Last Promotion - Years since the last promotion
- Years With Current Manager - Years with the current manager

Since the author had only access to the only data, it was impossible to determine which sampling technique was used to sample the respondents. It was not possible to determine the response rate and formular or statistical test used to determine the sample size. Most respondents were males (n=1764), they were married (n=1346) and did not work overtime (n=2108), they had life sciences qualifications (n=1212), and they were in sales execute positions (n=652).

3.3. Data analysis

The data were analysed in Python, a programming language used to analyse the data (McKinney, 2018). Data were analysed by means of frequencies, which are shown in section 1.2.2. The descriptive statistics comprised the mean (averages), median (middle), box plots and bar plots for visualisation. In addition, correlation analysis was conducted to show the relationship between predictors since, in Python, it only indicates non-categorical variables (McKinney, 2018). Bagging was used as an ensemble paradigm, and it is about evaluating the learner parallel and independently using bootstrap or re-sampling techniques. (McKinney, 2018). It is also said to be "black box estimators in random subsets of the original data set, then aggregate their predictions effectively to work out and construct the final prediction" (Subhashini & Gopinath, 2020, p.3333) Confusion matrix (refer to Figure 1) was used to evaluate the learners.

True negatives (TN): Cases in which prediction cases are 'No' and do not leave the company.	**False positives (FP):** Cases in which prediction cases are 'Yes', but do not leave the company. (type I error).
False negatives (FN): Cases in which prediction cases are 'No', but actually leave the company (type II error).	**True positives (TP):** Cases in which prediction cases are 'Yes' (employee will leave the company), and they do leave the company.

Figure 1: Attrition confusion matrix
Source: Khera and Divya (2019)

Various evaluation metrics and formulas were used to determine which learner outperformed others. The highest score in the training and test data were used to select the best-performing learner. In addition, the difference of 5 between training and test data to determine overfitting.

Table 2: Evaluation metric

Evaluation metric	Formula
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$
F1	$2*\frac{precision*recall}{precision+recall}$

Source: Author's own construction

3.4. Ethical consideration

Ethics there a moral way of conducting research, following certain principles (Zikmund, Babin, Carr & Griffin, 2013). The University of Texas Business at Austin Business School Data Analytics division granted the author permission to use the data analysed in this paper. In addition, the author adhered to the principle of anonymity and confidentiality by not disclosing the identity of the consulting company.

4. Results and Findings

The descriptive statistics, bivariate analysis and models that outperformed is discussed. The data in Table 2 showed that the youngest employee was 18. The mean and median of the daily and hourly rates are equal, suggesting that this variable was normally distributed. The monthly rate mean was slightly higher than the median. The data showed that the longest distance to travel to work was 29 kilometres. In terms of the Number of Companies Worked for, the mean -was higher than the median, suggesting that the data was slightly skewed to the right. The data showed that the percentage increase ranged from 11 to 25, and since the median is greater than the mean, this distribution shows negative skewness. It is observable that the mean of Total Working Years was higher than the median, suggesting that the data is skewed to the right. There was a difference in training opportunities, as some respondents had six opportunities and others had none. The range in the number of years and current role in the consulting company was 40 and 18, respectively. The range since the last promotion and with the current manager was 15 and 17 years, respectively.

Table 3: Descriptive statistics

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
Age	37	36	9	18	60
Daily Rate	802	802	403	102	1499
Hourly Rate	66	66	20	30	100
Monthly Rate	14313	14235	7116	2094	26999
Distance From Home	9	7	8	1	29
Number Companies Worked	3	2	2	0	9
Percentage salary increase	15	14	4	11	25
Total Working Years	11	10	8	0	40
Training Times Last Year	3	3	1	0	9
Years At Company	7	5	6	0	40
Years In Current Role	4	3	4	0	18
Years Since Last Promotion	2	3	0	0	15
Years With Current Manager	4	4	3	0	17

Source: Author's own construction

4.1 Bivariate analysis

In this section, the bivariate analysis is discussed. They are in the form of boxplots, stacked plots and correlations. The discussion focuses on the top features that predicted attrition the highest. The data showed that in Figure 2, the median of those who were least likely to churn is slightly higher than those who were least likely to. The data shows that the minimum year of those who are least likely to resign is 18 years, and at the 25th percentile, the age is 28 years. The data showed that the mean scores of those who earn on an hourly rate are likely to resign are the same who are least likely to resign. On average, employees who travel longer kilometres from home are more likely to attrite.

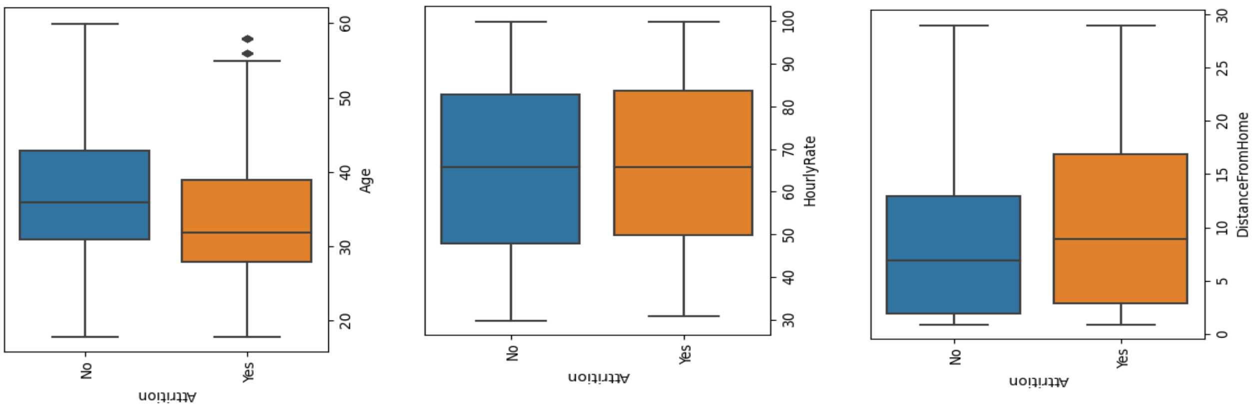


Figure 2: Age, hourly rate, distance from home and attrition
Source: Author's own construction

The data in Figure 3 showed that respondents with lower Daily rates and less monthly wages are likelier to attrite. It showed that respondents' monthly and hourly rates do not seem to impact employees' attritions, and a lesser salary hike also contributes to attrition.

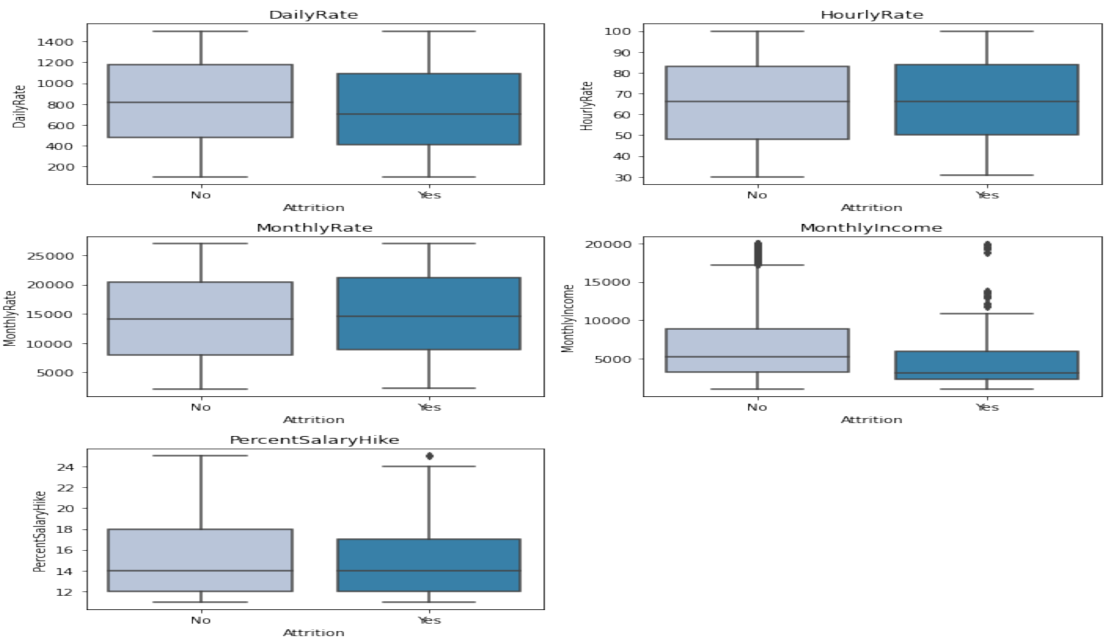


Figure 3: Daily, hourly and monthly rates, percent salary hike and monthly income
Source: Author's own construction

The data in Figure 4 showed that employees who worked for many companies and had less experience are most likely to attrite.

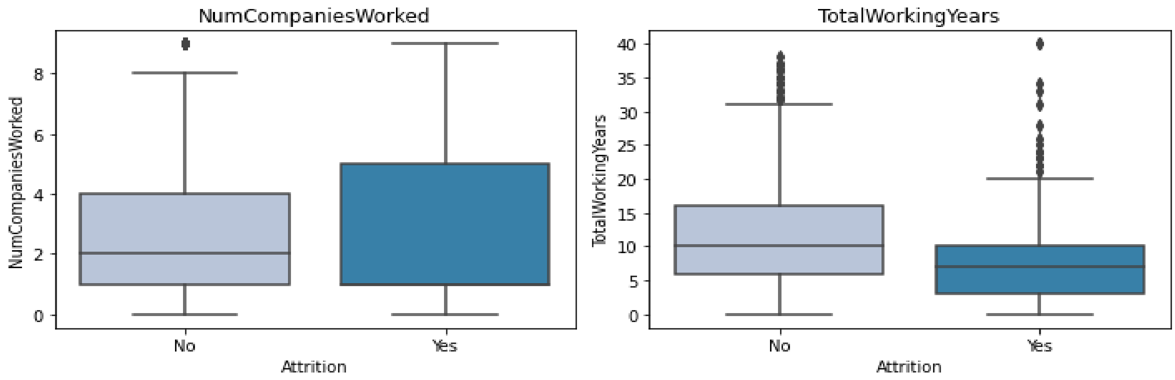
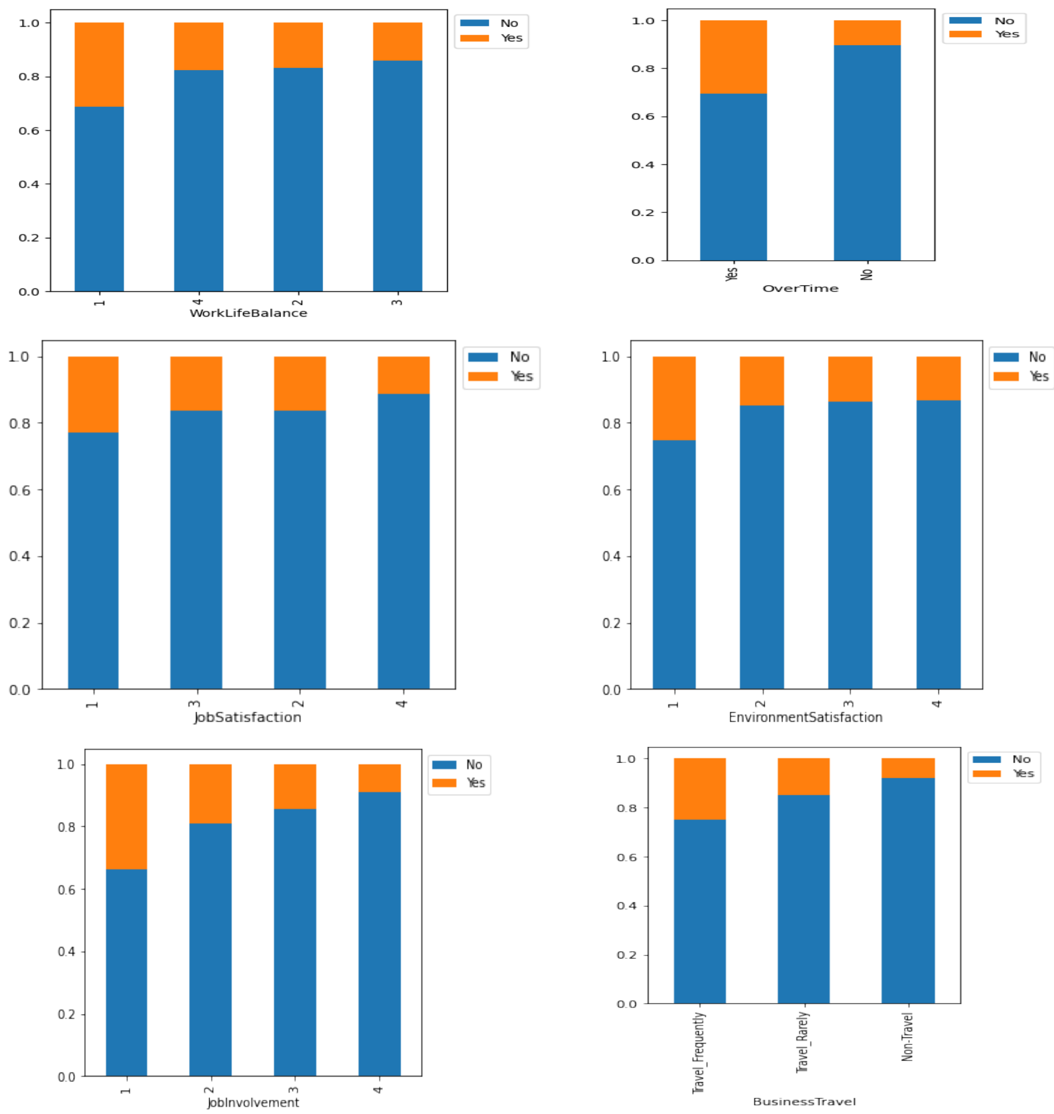


Figure 4: Number of companies worked, total working years and attrition
Source: Author's own construction

The data in Figure 5 showed that respondents who worked overtime, who rated job satisfaction, work-life balance the lowest and job involvement the lowest, who travelled frequently, who did not have stock options and are working as sales executives had the highest propensity to attrite.



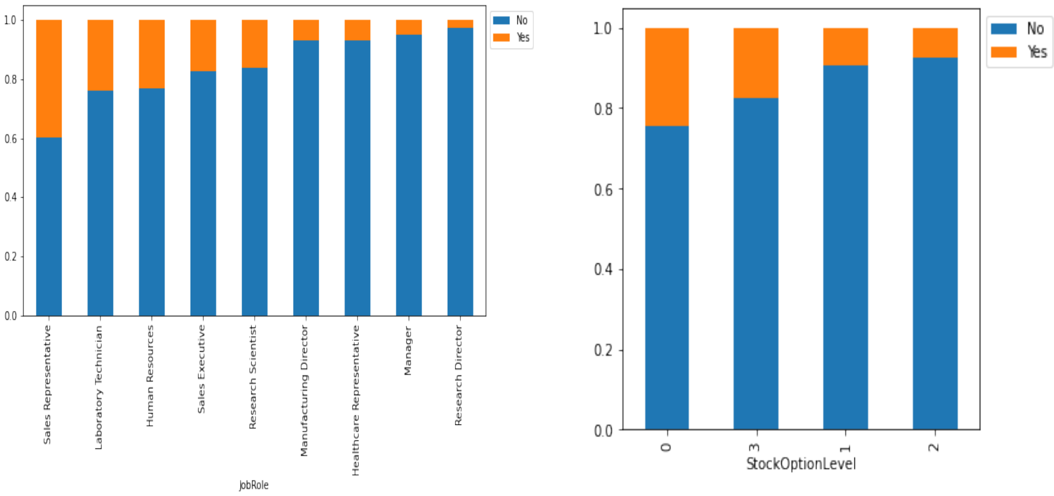


Figure 5: Work-life balance, overtime, job satisfaction, job involvement, job role, stock options and attrition
Source: Author's own construction

The data in Figure 6 showed that the highest correlation was between monthly income and job level (r=0.95). The second highest correlation was found between job level and total years of years (r=0.78). It can be observable that the relationship between performance rating and percentage salary hike (r=0.77). The same level of correlation is found between monthly income and the number of years.

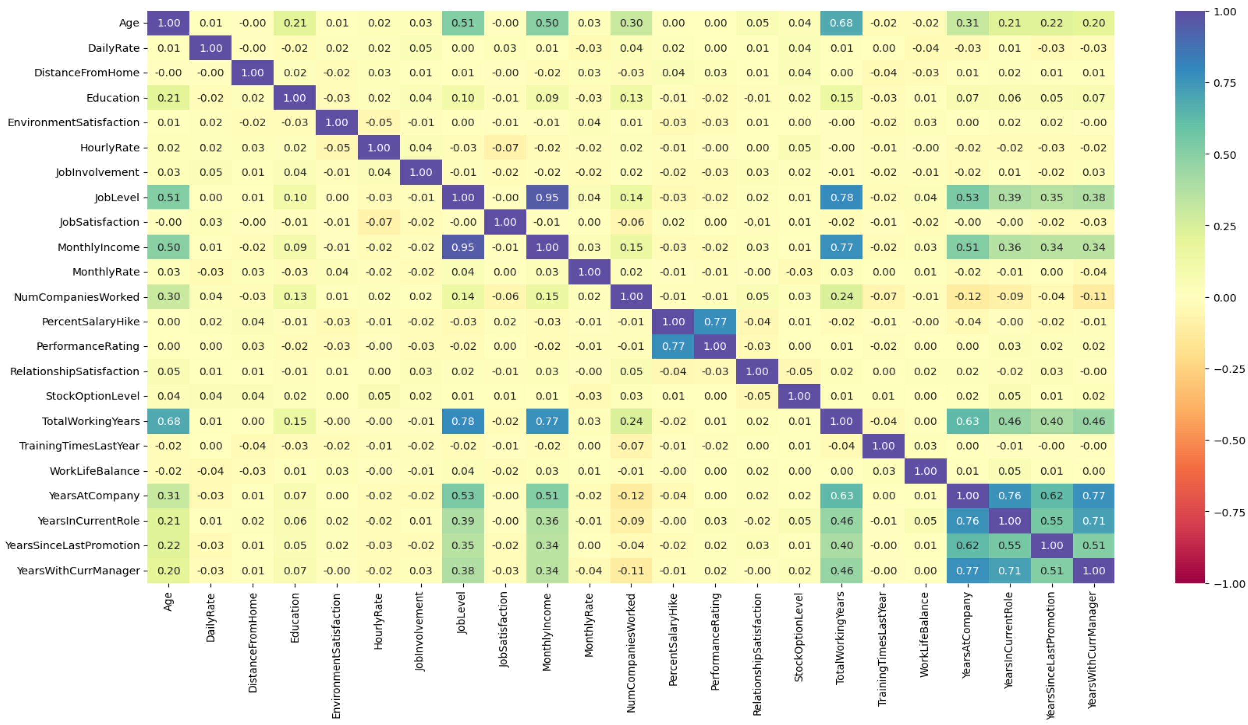


Figure 6: Relationships between predictors of attrition
Source: Author's own construction

Before building the model, outlier detection was conducted; as observed, most of the predictors did not have outlier issues. Other than detecting the outliers, additional data cleaning included dropping. Employee Number ID was neither a categorical nor a continuous variable.

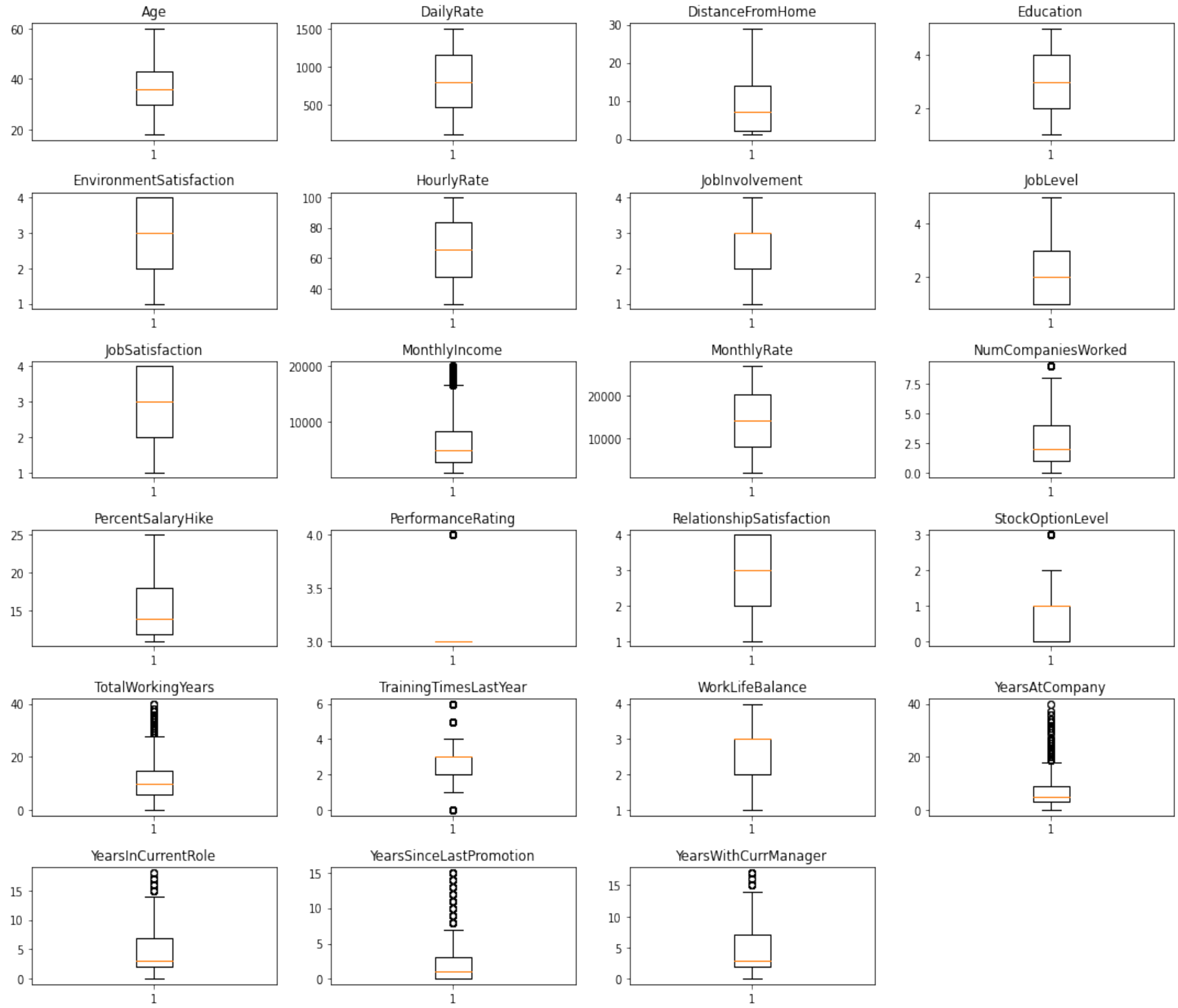


Figure 7: Outlier detection
Source: Author's own construction

4.2 Model building

Model building sing bagging (i.e. ensemble methods) requires experimenting with different models and choosing the one that performs the best or the best classifier. The classifiers used in this study are shown in Table 3. The first step of model building involves splitting the data into training and testing. The workflow training is shown in Figure 8, which shows that the training model is about training the model with noise and tuning it, and the testing model (Mahesh, 2018).

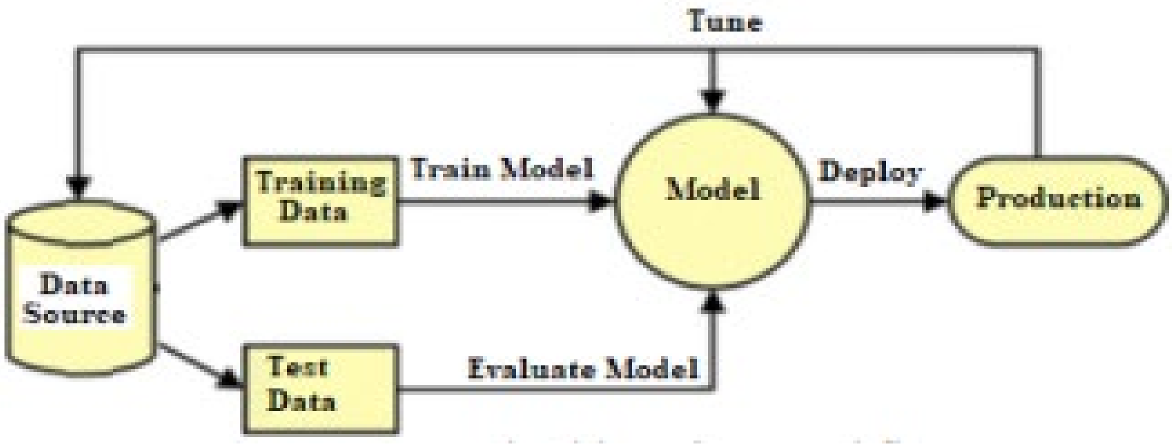


Figure 8: Training and testing
Source: Mahesh (2018, p. 3)

The splitting was 70 training and 30 test. After splitting the data, the former had rows (2058) and columns (44), and the latter had rows (882) and columns (44). The value counts were 83.90% for 0 (i.e. those who did not attrite) and 16.09% for 1 (i.e. those who did attrite). As it be observed from the data in Table 4, Random Forests Estimator (RFE) outperformed the other models.

Table 4: Comparison of all models

Training performance comparison								
Model Training	**DC**	**BC**	**WBC**	**RFC**	**WRFC**	**DTE**	**BE**	**RFE**
Accuracy	1.00	0.99	0.99	1.00	1.00	0.82	0.40	1.00
Recall	1.00	0.96	0.96	1.00	1.00	0.76	1.00	1.00
Precision	1.00	1.00	1.00	1.00	1.00	0.46	0.21	1.00
F1	1.00	0.98	0.98	1.00	1.00	0.58	0.35	1.00
Test performance comparison								
Model Training	**DC**	**BC**	**WBC**	**RFC**	**WRFC**	**DTE**	**BE**	**RFE**
Accuracy	0.94	0.95	0.91	0.96	0.96	0.79	0.37	0.96
Recall	0.83	0.70	0.70	0.80	0.78	0.64	0.95	0.80
Precision	0.79	0.94	0.92	0.90	0.96	0.40	0.20	0.95
F1	0.81	0.81	0.80	0.87	0.87	0.87	0.33	0.87

Note: DC stands for Decision Tree, BC stands Bagging Classifier, WBC stands Weighted Bagging Classifier, RFC stands for Random Forest Classifier, WRFC stands for Weighted Random Forest Classifier, DTE stands for Decision Tree Estimator, BE stands for Bagging Estimator and stands for Random Forest Estimator
Source: Author's own construction

The RFE training and testing confusion matrix is shown in Figure 9 after experimentation.

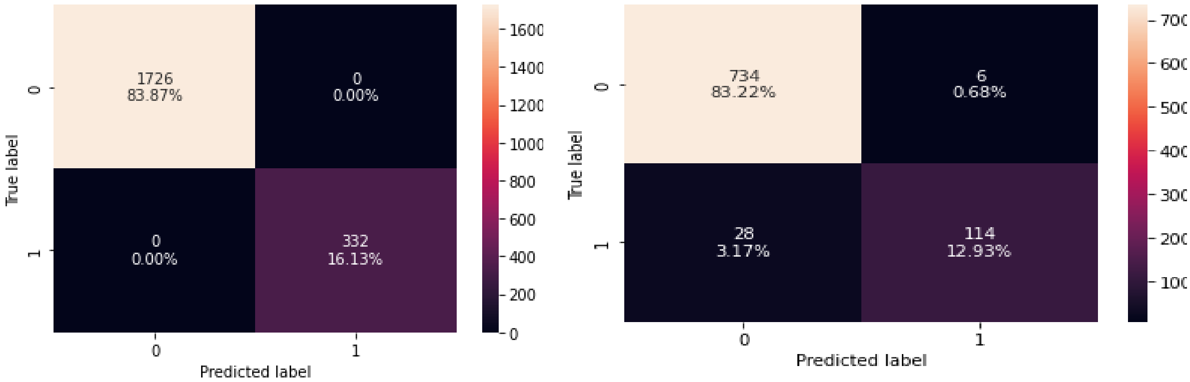


Figure 9: RFE training and testing confusion matrices
Source: Author's own construction

The feature importance of RFE was shown in Figure 10. The data show that 4 out 8 important features/predictors that can make employees attrite were related or variants of compensation. The highest predictor was monthly income, the second highest was working overtime, and the third highest was daily rate. The data showed that performance rating, the manager and the job role of the director were the minor predictors of employee attrition.

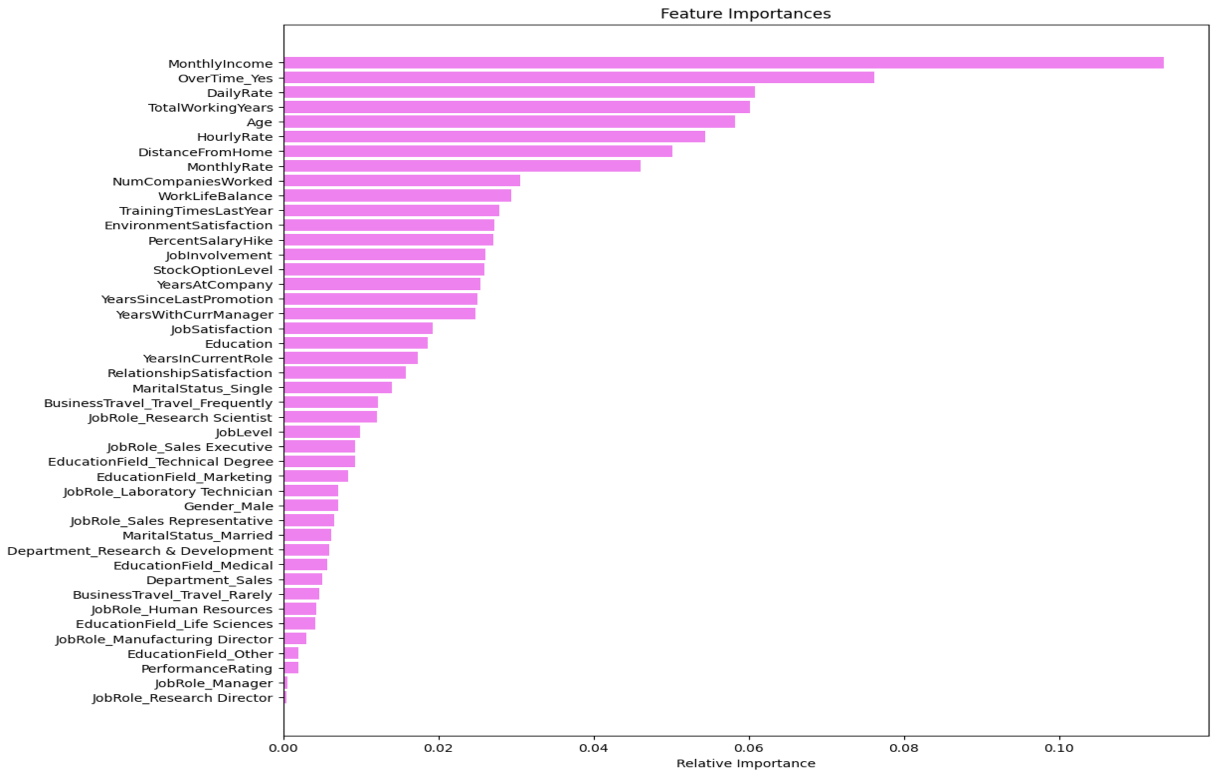


Figure 10: Feature importance of RFE
Source: Author's own construction

5. Managerial Implications

This study has created insights for managers to develop strategies to diminish attrition. Since variants of income (i.e. hourly, daily and monthly), it suggested managers should benchmark their salary structure and adjust according to the market. This means that managers should strive to develop strategies that would make them lead the market. In addition, managers should build flexibility and travelling policy. Such a policy should be developed or revised with employees, as it has been found that married employees with younger children are most likely to attrite (Adriano & Callaghan, 2020). The latter will ensure that it balances the business travel intervals amongst the staff. In addition, managers develop a work-from-home policy, and this assists them in retaining Generation Z, who prioritise flexibility (Benito-Osorio et al., 2014).

Since working for many organisations contributed to attrition, such information should be used as a determinant to hiring employees during recruitment and selection. Of this, managers should be cautious when hiring employees with such profiles. If they possess scarce skills, managers can offer stock options in the organisation. These stock options should have conditions like when employees attrite within 4 years, and they would lose their shares in the organisations. Managers should engage employees working as sales executives so that they can develop retention strategies for them. Lastly, after conducting performance appraisals, it is recommended that managers should send employees with performance gaps for training.

6. Conclusions, Limitations and Future Research

The paper aimed to identify important features that predict employee attrition at the multination consulting corporation. Unlike other HRM studies that used SEM and few predictors (Schaap & Olckers, 2020), this study used an ensemble technique to identify which essential predictors of employee attrition. Following this approach was already advocated by Oswald et al. (2019), who encouraged HRM and I/O scholars to explore another methodological approach when investigating topics like attrition, for example. In addition, the advantage of using ensemble methods is that “human resource problems, unlike physical systems, cannot be defined by a scientific-analytical formula” (Subhashini & Gopinath, 2020).

Similar to Fakhri et al. (2020) and Mobley et al. (1979), in this study, it was found that 4 out 8 top important predictors of attrition were remunerated related. Those remunerated related various remuneration-related variables were monthly income, daily, hourly and monthly rates. This finding can be attributed to the consulting company not paying in at the median and above the median in the consulting labour market. Paying employees above the labour make has been the best retention strategy to mitigate employee attrition (Bussin, 2018). Mowday (1984) is of the view that retention should only focus on talented and high-performing employees. The same author believes that when non-performing employees attrite, it is suitable for organisations that have created a strategy hiring access labour that trained them and have taught a culture organisation committed. In such an organisation, there will be no need to incur acquisition, onboarding and training costs which have been costed around by the Society of HRM to between 30% to 40% for employees at the entry-level, 125% at middle management and 200% at senior management (Gross, 2022).

In addition, in line with Mobley et al. (1979) framework, this study found that labour economic factors like age and working years contribute to employees' attrition. Such employees are the ones who have less working experience and might be young. Benito-Osorio et al. (2014) established Generation Z (i.e. the late 1990s to early 2000s), who resigned at the highest frequencies of attrition compared to the older employees. The same authors recommended that organisations implement a flexible schedule for Generation Z as they value it the most rather than paying them in labour with the labour market.

Working distance from working work was also found to contribute to employees' attrition. This can be attributed to employees' paying travelling costs and spending long hours in traffic. The unintended consequence of employees spending long hours is that they are tired when they arrive at the workplace. Since they arrive at work tired, they cannot perform optimally and are rated poorly by managers (Rodriguez, Swain & Springer, 2020). This makes the organisation not achieve its strategic objectives and lose its competitive edge (Liu & Raghuram, 2021). Notably, in this study, performance rating is not critical, which suggests that employees are satisfied with this HRM process.

Another significant predictors of employee attrition was the number of companies worked and work-life balance. In terms of the latter, it was found working long hours had an adverse effect, and employees with low work-life balance who are married and have younger children are most likely to attrite (Adriano & Callaghan, 2020). A plausible reason for employees resigning from one consulting company to another is that knowledge workers are mobile, and their skills are in demand (Sheidaee, Philsoophian & Akhavan, 2022).

Similar to the literature (Fakhri et al., 2020), it was found that training was also a predictor of employee attrition. Another finding was that environmental satisfaction plays are role in employee attrition. It is defined as employees' subjective feelings about the building they work at (Shin, 2016). The data showed that environmental job satisfaction correlated the lowest with other predictors; the highest correction was with monthly income (r=0.04). The highest correlation between predictors was between monthly income and job level (r=0.95) and the total number of years (r=0.78). This can be interpreted when employees' job grades and experience contribute to higher wages.

Despite this study making a contribution to the HRM body of knowledge by including environmental job satisfaction and using different attrition matrices unlike other South African studies (Dhanpat & de

Braine, 2022; Schaap & Olckers, 2020), it had limitations. One of its limitations is that it used a cross-sectional research design, which painted a once picture of the research topic. Based on this, it is recommended that a longitudinal study be conducted in the future. Since it was quantitative, it could determine in-depth reasons why certain employees were not offered training and why they rated job satisfaction and work-life balance low. This can be established by conducting a qualitative study. Since this study is one of fewer in HRM literature, it's recommended that a similar approach should be followed by researchers who are not focussing on confirming Mobley et al. (1979) attrition framework in their context using few predictors.

In conclusion, this study used ensemble methods to identify which features predicted attrition in consulting organisation. It was influenced by Mobley et al. (1979) attrition framework and used 32 predictors. The model comparison showed that RFE outperformed other models. It be concluded that these were the top 5 predictors of attrition: monthly income, overtime, daily rate, total working years and age.

Acknowledgments

The author's is grateful to the DVC of Teaching and Learning for paying for the author's post graduate data science diploma at Texas Business at Austin Business School Data, where was given permission to use the data analysed in this study.

REFERENCES

Adriano, J. & Callaghan, C.W., (2020). Work-life balance, job satisfaction and retention: Turnover intentions of professionals in part-time study. *South African Journal of Economic and Management Sciences*, 23(1), a3028. https://doi.org/10.4102/sajems.v23i1.3028.

Akgunduz, Y., Gök, Ö.A. & Alkan, C. (2020). The effects of rewards and proactive personality on turnover intentions and meaning of work in hotel businesses. *Tourism and Hospitality Research*, 20(2),170–183.

Al Mamun, C.A. & Hasan, M. (2017). Factors affecting employee turnover and sound retention strategies in business organisation: a conceptual view. *Problems and Perspectives in Management,* 15(1), 63-71.

Barkhuizen, N., Lesenyeho, D. & Schutte, N. (2020). Talent retention of Academic Staff in South African Higher Education Institutions. International Journal of Business and Management Studies, 12(1), 177-190.

Behravesha, E., Tanovaa, C. & d Abubakar, A.M. (2020). Do high-performance work systems always help to retain employees or is there a dark side? *The Service Industries Journal*, 40(11), 825-845.

Benito-Osorio, D., Muñoz-Aguado, L.& Villar, C. (2014). The Impact of Family and Work-Life Balance Policies on the Performance of Spanish Listed Companies. 17(4): 214-236.

Bussin, M . (2018). Retention strategies: The key to attract and retain excellent employees. Randburg: KR Publishing.

Castro, M.L. & Martins, N. (2010). The relationship between organisational climate and employee satisfaction in a South African information and technology organisation. *SA Journal Industrial Psychology*, 36: 1-9.

Cotton, J.L. & Tuttle, J.M. (1986). Employee Turnover: A Meta-Analysis and Review with Implication for Research. The Academy of Management Review, 11((1), 55-70.

Dhanpat, N. & de Braine, R. (2022). The Influence of Retention Factors on Nurses' Turnover Intention. IBC Proceedings.

Dutta, D. & Bandyopadhyay, S.K. (2020). Employee attrition prediction using neural network cross validation method. International Journal of Commerce and Management Research, 6(3), 80-85.

Enrich. (2023). The Cost of Replacing an Employee and the Role of Financial Wellness. Available at: https://www.enrich.org/blog/The-true-cost-of-employee-turnover-financial-wellness-enrich#:~:text=Research%20by%20SHRM%20suggests%20that,overall%20losses%20to%20the%20company.

Fakhri, M., Nurnida, I., Winarno, A. & Suryana, D. (2020). Characteristics of Quality of Work Life on Employees at Consultant Company in Indonesia. *Journal of Asian Finance, Economics and Business*, 7(11), 1105- 1111.

Fallucchi, F., Coladangelo, M., Giuliano. R. & William De Luca, E. (2020). Predicting Employee Attrition Using Machine Learning Techniques. *Computers*, 9(86), 1-17.

Grobler, P.A., Wärnich, S., Carell, M.R., Elbert, R.D. & Hatlield (2011). Human Resource Management in South Africa. (4th.). United Kingdom. Cengage Learning.

Grobler, A. & Jansen van Rensburg, M. (2019). Organisational climate, person–organisation fit and turn over intention: a generational perspective within a South African Higher Education Institution. *Studies in Higher Education*, 44:11,

Gross, A. (2022). Think a Company Retirement Plan is Expensive? Try Not Having. Available at: https://www.ijoinsuccess.com/think-a-company-retirement-plan-is-expensive-try-not-having-one/

Hancock, J.I., Allen, D.G., Bosco, F.A., McDaniel, K.R. & Pierce, C.A. (2013). Meta-Analytic Review of Employee Turnover as a Predictor of Firm Performance. *Journal of Management*, 39(3), 573-603.

Hom, P.W., Lee, T.W., Shaw, J.D. & Haushnecht, J.P. (2017). One Hundred Years of Employee Turnover Theory and Research. *Journal of Applied Psychology*, 102(3), 530–545.

Hughes, M.G., O'Brien, E.L. Reeder, M.C. & Purl, J. (2020). Attrition and reenlistment in the Army: Using the Tailored Adaptive Personality Assessment System (TAPAS) to improve retention. *Military Psychology*, 32(1), 36-50.

Khera, S.N. & Divya, N. (2019). Predictive Modelling of Employee Turnover in Indian IT Industry Using Machine Learning Techniques. Vision, 23(1), 12–21

Liu, X. & Raghuram, S. (2021). The effects of latent withdrawal profiles on employee turnover, destinations and job performance. *Human Resource Management Journal,* 32,384–405.

Mahesh, B. (2018). Machine Learning Algorithms - A Review. *International Journal of Science and Research,* 9(1), 391- 386.

Maleka, M. (2012). An in-depth investigation of the factors contributing to employee dissatisfaction at the Business Application Solution Centre (BASC), Eskom. D Litt et Phil thesis. University of South Africa.

Maleka, M.J. (2018). The biographical and human resource management predictors of union membership engagement of low- and middle-income workers. Journal of economics and behavioral sciences 10(1):207-216

Mathieu, C., Fabi, B., Lacoursière, B. & Raymond, R. (2016). The role of supervisory behavior, job satisfaction and organisational commitment on employee turnover. Journal of Management & Organization, 22(1), 113–129.

McKinney, W. (2018). Python for Data Analysis Data Wrangling with Pandas, NumPy and IPython. (2 nd.). O'Reilly Media, Inc: United States.

Mello, J.A. (2015). Strategic human resource management. (4th ed.). Australia: Cengage Learning.

Memon, M.A., Salleh, R., Mirza, M.Z., Cheah, J.-H., Ting, H., Ahmad, M.S. & Tariq, A. (2021). Satisfaction matters: the relationships between HRM practices, work engagement and turnover intention. International *Journal of Manpowe*r, 42(1), 21-50. https://doi.org/10.1108/IJM-04-2018-0127

Michaels, C. E. & Spector, P.E. (1982). Causes of Employee Turnover: A Test of the Mobley, Griffeth, Hand, and Meglino Model. *Journal of Applied Psychology*, 67(1),53-59.

Mobley, W.H., Griffeth, R.W., Hand, H. H. & Meglino, B.M. (1979). Review and Conceptual Analysis of the Employee Turnover Process. *Psychological Bulletin*, 86(3), 493-522.

Mowday, R.T. (1984). Strategies for Adapting to High Rates of Employee Turnover. *Human Resource Management*, 23(4), 365-380.

Nouri, H. & Parker, R.J. (2019). Turnover in public accounting firms: a literature review. *Managerial Auditing Journal*, 35(2), 294-321.

Oswald, F.L., Behrend, T.S., Putka, D.J. & Sinar, E. (2019). Big Data in Industrial-Organizational Psychology and Human Resource Management: Forward Progress for Organizational Research and Practice. Annual Review of Organizational Psychology and Organizational Behavior. Available from: https://doi.org/10.1146/annurev-orgpsych-032117- 104553

Pratt, M., Boudhane, M. & Cakula, S. (2021). Employee Attrition Estimation Using Random Forest Algorithm. Baltic J. *Modern Computing*, 9(1), 49-66.

Porter, L. W., & Steers, R. M. (1973). Organisational, work, and personal factors in employee turnover and absenteeism. *Psychological Bulletin,* 80(2), 151–176. https://doi.org/10.1037/h0034829

Punnoose, R. & Ajit, P. (2016). Prediction of Employee Turnover in Organizations using Machine Learning Algorithms. A case for Extreme Gradient Boosting. *International Journal of Advanced Research in Artificial Intelligence,* 5(9), 22-26.

Rodriguez, L.A., Swain, W.A. & Springer, M.G. (2020). Sorting Through Performance Evaluations: The Influence of Performance Evaluation Reform on Teacher Attrition and Mobility. *American Educational Research Journal*, 57(6) 2339–2377.

Saunders, M., Lewis, P. & Thornhill, A. (2019). Research methods for business students. (8th ed.). Harlow: Pearson.

Sheidaee, S., Philsoophian, M. & Akhavan, P. (2022). The effect of intra-organizational knowledge hiding on employee turnover intentions: the mediating role of organisational embeddedness: a case study of knowledge workers of IRIB. *Journal of Organizational Effectiveness: People and Performance*, 9 (3) 422-44.

Shin, J. (2016). Toward a theory of environmental satisfaction and human comfort: A process-oriented and contextually sensitive theoretical framework. *Journal of Environmental Psychology*, 45, 11 – 21.

Skosana, T.B., Maleka, M.J. & Ngonyama-Ndou, T. 2021. Predictors of affective commitment at municipalities in the Nkangala District, Mpumalanga. SA Journal of Human Resource Management. *SA Journal of Human Resource Management*/SA Tydskrif vir Menslikehulpbronbestuur,19(0), a1567. https://sajhrm.co.za/index.php/sajhrm/article/view/1567

Spector, P.E. (1997). Job satisfaction: Application, assessment, causes and consequences. Thousand Oaks, CA: Sage.

Subhashini, M. & Gopinath, R. (2020). Employee Attrition Prediction in Industry Using Machine Learning Techniques. *International Journal of Advanced Research in Engineering and Technology,* 11(12), 3329-3341.

Tirta, A.H. & Enrika, A. (2020). Understanding the impact of reward and recognition, work life balance, on employee retention with job satisfaction as mediating variable on millennials in Indonesia. *Journal of Business and Retail Management Research*, (JBRMR), 14 (3), 88-98.

Tshwane, G.S., Maleka, M.J., & Tladi, P.M. 2023. Investigating turnover intention in a financial organisation in Gauteng. SA *Journal of Human Resource Management/SA Tydskrif vir Menslikehulpbronbestuur*, 21(0), a2177. https://doi. org/10.4102/sajhrm. v21i0.2177

Winarno, A., Prasetio, A.P., Luturlean, B.S., & Wardhani, S.K. (2022). The link between perceived human resource practices, perceived organisational support and employee engagement: A mediation model for turnover intention. *SA Journal of Human Resource Management/SA Tydskrif vir Menslikehulpbronbestuur,* 20(0), a1802. https://doi.org/10.4102/sajhrm.v20i0.1802

Zikmund, W., Babin, B., Carr, J. & Griffin, M. (2013). Business research methods. (9th ed.). Mason, OH: South-Western.

Zounemat-Kermani, M., Batelaan, O., Fadaee, M. & Hinkelmann, R. (2021). Ensemble machine learning paradigms in hydrology: A review. *Journal of Hydrology*, 598. https://doi.org/10.1016/j.jhydrol.2021.126266.