Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 4 Issue 2 (2024)

The Role of Demographic and Behavioural Data in Predictive Analytics for Employee Retention

Mr. Debabrata Sahoo[*1], Dr. Smaraki Pattanayak[2], Dr. Phalgu Niranjana[3]
[*1]Doctoral Research Scholar, School of Business, ASBM University, Bhubaneswar, Odisha, INDIA. E-mail ID: debabrata.sahoo1612@gmail.com
[2]Associate Professor, School of Business, ASBM University, Bhubaneswar, Odisha, INDIA. E-mail ID: smaraki.pattanayak@asbm.ac.in
[3]Professor, School of Business, ASBM University, Bhubaneswar, Odisha, INDIA. E-mail ID: phalgu.niranjana@asbm.ac.in

***Corresponding author: Mr. Debabrata Sahoo**
***E-mail ID: debabrata.sahoo1612@gmail.com**

ABSTRACT
Organisations continue to face a significant difficulty in employee retention, which calls for creative solutions to anticipate and reduce turnover. The purpose of this study is to better understand how behavioural and demographic data might improve predictive analytics for employee retention. A specific machine learning algorithm was deployed, namely Gradient Boosting Technique to develop a predictive model which was then assessed by utilising an extensive dataset from a mid-sized technology company, which included behavioural indicators (performance ratings, attendance records, engagement survey scores, and training participation) and demographic variables (age, gender, education level, marital status, and tenure). The results show that the accuracy of turnover estimates is greatly increased by integrating behavioural and demographic data. The most important indicators of attrition were found to be behavioural elements, specifically performance and engagement levels, however demographic factors like age and tenure also had a big impact. This study emphasises the value of a comprehensive approach to predictive analytics in HR, which helps businesses to develop focused retention plans that increase employee stability and productivity. In order to improve retention results and further develop prediction models, future research should investigate other data sources.

Keywords: Predictive Analytics, Gradient Boosting Technique, Machine Learning Algorithm, Employee Retention, Demographic Data, Behavioural Data

1. INTRODUCTION

Employers in a variety of sectors are very concerned about employee retention. Excessive staff turnover rates can result in significant expenses, such as those related to hiring and onboarding new hires, institutional knowledge loss, and low morale among current employees. Human resource (HR) professionals now prioritise understanding and forecasting employee turnover as a result of these problems. Conventional methods of handling employee turnover, such satisfaction surveys and departure interviews, frequently yield little information and are reactive rather than proactive. As a result, predictive analytics has become a potent instrument for anticipating and reducing staff turnover. By using previous data to predict future occurrences, predictive analytics helps organisations identify high-risk personnel and put tailored retention tactics in place. Predictive models can find patterns and connections by combining several data sets that are not immediately visible using traditional techniques. While there are many different kinds of data accessible, behavioural and demographic data have shown to be very useful in forecasting employee turnover.

Aspects like age, gender, tenure, marital status, and education level are examples of demographic data. These variables give a static picture of the conditions and background of an employee, revealing possible risk factors for turnover. For example, compared to their older or more tenured peers, younger employees or those with less tenure may have different career goals and levels of job satisfaction. Organisations may customise their retention tactics to the unique requirements and preferences of various staff groups by having a thorough understanding of these demographic tendencies. On the other hand, dynamic information about an employee's activities and interactions inside the company is included in behavioural data. This comprises engagement levels, attendance records, performance metrics, involvement in training, and other measures of an employee's regular conduct and involvement. Behavioural data is an essential part of predictive models since it offers a real-time perspective of an employee's pleasure and engagement.

Predictive analytics, which incorporates behavioural and demographic data, provides a thorough method for comprehending and forecasting employee attrition. While demographic information may be used to divide up the workforce and pinpoint general risk indicators, behavioural information offers a more detailed picture of the experiences

of specific employees and possible red flags of attrition. Organisations may create prediction models that are more accurate and useful by merging several data sources.

The purpose of this study is to investigate how behavioural and demographic data relate to predictive analytics for employee retention. It will specifically look at the major behavioural and demographic variables that affect employee turnover and assess how they affect prediction model accuracy. The goal of the study is to give HR professionals useful information through this analysis so they may improve retention methods and lower turnover rates.

The paper will be organised as follows: the literature review will include a summary of previous studies on predictive analytics in HR, emphasising the use of behavioural and demographic data. The steps involved in gathering data, preparing it, creating and validating prediction models, and more are covered in the methodology section. The study's conclusions, including the effectiveness of various models and the importance of various predictors, will be presented in the results section. Ultimately, the talk will evaluate the findings, go over how they affect HR procedures, and offer ideas for further study. The goal of this thorough investigation is to add to the expanding body of knowledge on predictive analytics in HR and offer useful recommendations for enhancing employee retention.

2. LITERATURE REVIEW

The desire to better understand and control employee turnover has led to a major increase in the use of predictive analytics in human resources (HR). This review of literature looks at previous studies on predictive analytics in HR, with a particular emphasis on how behavioural and demographic data might be used to predict employee retention.

2.1 Predictive Analytics in HR

Statistical methods and machine learning algorithms are used in predictive analytics to examine past data and forecast future results. Predictive analytics is used in the HR domain to forecast employee behaviours like as performance, engagement, and turnover. Predictive analytics has the ability to revolutionise HR procedures by offering data-driven insights that assist in making strategic decisions, as several studies have demonstrated.

For example, Bassi (2011) highlights how predictive analytics may assist companies in seeing patterns and trends in worker conduct that are not immediately visible when using conventional techniques. With the use of this capacity, HR professionals may transition from a reactive to a proactive strategy, recognising and resolving problems before they become more serious.

2.2 Demographic Data

Aspects including age, gender, education level, marital status, and tenure are all included in the category of demographic data. These variables give an overview of an employee's background and are comparatively static. Studies have indicated that employee turnover may be considerably impacted by demographic characteristics. Significant relationships between demographic characteristics and socioeconomic results are found in a thorough assessment of the literature, underscoring the significance of demographic study in a variety of sectors.

2.3 Age and Tenure

Higher turnover rates are typically observed among younger personnel and those with shorter tenure. Younger workers are frequently in the early stages of their careers and may leave companies in quest of greater chances, according to Mitchell *et al.* (2001). In a similar vein, Griffeth *et al.* (2000) contend that shorter-tenured staff members are more likely to depart since they have not yet formed deep ties to their companies.

2.4 Gender

The impact of gender on turnover is more nuanced. Some studies, such as those by Hom *et al.* (1992), suggest that women may have higher turnover rates due to various factors, including work-life balance issues and organizational support. However, other studies indicate that the turnover gap between men and women is narrowing as organizations implement more inclusive policies.

2.5 Education Level

Turnover is also influenced by education level. Higher educated workers could have more options in the labour market, which would make them more likely to leave for better roles. But according to Koster *et al.* (2011), highly educated workers may also report better job satisfaction if their responsibilities fit their talents, which might lower turnover.

2.6 Behavioral Data

Information on the activities and interactions of employees inside the company is included in behavioural data. This kind of data is dynamic and offers information about the performance, pleasure, and engagement of a worker. Furthermore,

processing and interpreting complicated behavioural information has become much easier because of developments in machine learning and data analytics, which provide deeper insights and useful knowledge.

2.7 Performance Metrics

Performance measures, such productivity ratings and performance reviews, are important markers of worker satisfaction and possible attrition. If high-performing workers are not given enough credit or opportunity for growth, they may get unsatisfied and quit. On the other hand, poor performance might be a sign of disengagement, which happens before turnover (Allen *et al.,* 2010).

2.8 Attendance Records

Attendance records are important behavioural markers as well, including absenteeism and punctuality. Regular absences may be an indication of work unhappiness or disengagement (Markham *et al.,* 2002). Employers may use this information to spot workers who are in danger of quitting and take action before they do.

2.9 Engagement Levels

Worker engagement and turnover are highly correlated. Disengaged workers are more likely to quit their jobs, whereas engaged workers are more likely to stick with their companies. Low turnover intentions are correlated with high engagement levels, according to Schaufeli and Bakker's (2004) research. Workplace interactions, involvement in corporate events, and surveys are all ways to gauge employee engagement.

2.10 Training Participation

Employees that take part in training programmes demonstrate their dedication to their own and their careers' growth. According to Ragins *et al.* (1999), workers who participate in training are less likely to quit because they feel more involved in the company and recognise their potential for advancement. Effective training participation ensures that the employees are up-to-date with industry trends, and technological advancements, thereby improving overall organizational competitiveness.

2.11 Integration of Demographic and Behavioral Data

Predictive analytics's combination of behavioural and demographic data offers a comprehensive understanding of the variables affecting employee attrition. Organisations may create more precise prediction models and customise retention tactics to meet the unique requirements of certain employee groups by integrating these data types. In modern research, the integration of behavioural and demographic data has become a potent technique that combines conventional demographic analysis with insights from people's activities and preferences. This synthesis provides a thorough understanding of populations, allowing for more sophisticated and successful approaches in a range of contexts.

2.12 Predictive Model Development

Predictive models that combine behavioural and demographic data have proven to be useful in a number of studies. For instance, Kaur and Fink (2017) created a prediction model that used behavioural and demographic factors to accurately identify employees who were at-risk. According to their research, models that used both forms of data fared better than those that used only one.

2.13 Feature Significance Analysis

The most important turnover predictors may be found with the use of feature significance analysis. Evaluating each variable's contribution to the prediction model is part of this procedure. According to research by Jain and Gautam (2014), behavioural data—specifically, engagement and performance scores—were some of the most important indicators of employee turnover. Age and other demographic variables, including tenure, were also quite important.

2.14 Ethical Considerations

Predictive analytics in HR brings up a number of ethical issues. The quantity of data collection and analysis may cause discomfort for employees, making privacy considerations of utmost importance. Employers need to make sure that employees give their informed permission and that data gathering is transparent. Predictive algorithms also run the danger of bias, which might result in the unjust treatment of particular workforce groups. In order to reduce these hazards, researchers such as Raghavan *et al.* (2020) stress the need of ethical standards and procedures.

3. RESEARCH METHODOLOGY

3.1 Sample and Data Collection

Data was gathered over a five-year period from 2019 to 2023 from 50 randomly selected employees of a mid-sized technology business in order to examine the role of behavioral and demographic data in predictive analytics for employee retention. The comprehensive dataset comprised behavioral indicators such as performance evaluations, attendance records, engagement survey scores, and training participation, along with demographic characteristics including age,

gender, education level, marital status, and tenure. By integrating these diverse data points, the study aims to identify the predictive patterns using machine algorithm that could inform strategies for improving employee retention and fostering a more engaged and stable workforce.

Table 3.1: Sample Dataset of the Mid-Sized Technological Company

Year	Age (in years)	Gender	Educational Level	Marital Status	Tenure (in years)	Performance Rating	Attendance Record	Engagement Score	Training Participation
2019	29	Female	Bachelors	Married	2	5	207	8	3
	44	Female	Diploma	Widowed	7	2	358	6	6
	47	Male	Masters	Married	9	4	229	10	3
	27	Male	Masters	Single	1	4	360	9	1
	36	Female	Diploma	Married	6	4	270	7	2
	30	Male	Diploma	Single	5	3	177	10	4
	55	Female	Bachelors	Married	14	2	336	5	23
	59	Male	Masters	Married	17	3	345	8	17
	30	Female	Diploma	Single	3	5	296	2	5
	30	Female	Diploma	Married	3	2	287	7	7
2020	58	Female	Masters	Single	16	1	260	4	46
	36	Female	Diploma	Divorced	7	4	238	6	14
	41	Male	Diploma	Married	11	5	213	7	37
	34	Male	Masters	Married	9	1	184	10	26
	43	Female	Bachelors	Divorced	2	1	274	4	4
	49	Male	Bachelors	Married	1	4	163	2	2
	48	Male	Diploma	Married	4	1	355	4	3
	35	Male	Masters	Single	3	4	172	7	5
	53	Male	Diploma	Married	6	5	197	2	6
	57	Female	Diploma	Married	12	1	257	4	19
2021	55	Female	Diploma	Married	10	4	187	9	31
	36	Female	Masters	Single	7	4	235	2	11
	57	Male	Bachelors	Divorced	11	4	200	6	14
	44	Female	Bachelors	Divorced	9	2	218	5	16
	38	Male	Masters	Single	2	4	248	8	3
	32	Male	Diploma	Single	2	4	226	4	4
	55	Male	PhD	Married	13	1	348	1	9
	31	Female	Diploma	Widowed	3	3	215	10	7
	49	Male	PhD	Married	6	2	150	3	2
	37	Male	Bachelors	Married	7	4	158	7	3
2022	59	Female	PhD	Married	21	4	345	8	42
	51	Female	Bachelors	Married	19	1	242	4	23
	25	Female	Diploma	Single	1	3	283	6	0
	44	Male	Masters	Married	4	5	356	5	3
	59	Female	Diploma	Married	18	3	267	7	18
	36	Female	Diploma	Married	7	2	156	2	4
	59	Male	Bachelors	Married	12	3	282	4	17
	28	Male	Diploma	Single	2	4	185	10	6
	44	Female	Masters	Married	6	2	210	4	11
	34	Female	Diploma	Divorced	8	3	288	7	13
2023	41	Male	Diploma	Divorced	9	5	273	4	7
	25	Female	Bachelors	Single	1	4	165	10	1
	30	Male	Masters	Divorced	3	2	324	6	2
	45	Female	Diploma	Married	5	2	233	9	6
	35	Female	Bachelors	Married	6	4	189	1	9

	55	Female	Bachelors	Married	13	2	263	10	14
	52	Female	PhD	Widowed	11	2	165	8	19
	56	Male	Diploma	Married	17	4	225	10	31
	26	Male	Masters	Single	1	4	181	7	0
	60	Male	PhD	Married	16	4	336	9	29

3.2 Columns Description

- **Year:** The year of data collection
- **Age:** Age of the employee (in years, as on 31 December 2023)
- **Gender:** Gender of the employee (Male/Female)
- **Educational Level:** Highest level of education achieved (Bachelors/Diploma/Masters/PhD)
- **Marital Status:** Marital status of the employee (Single/Married/Divorced/Widowed)
- **Tenure:** Number of years the employee has been with the company
- **Performance Rating:** Employee performance rating on a scale of 1 to 5 in the last quarter of the year 2023 (1-Lowest and 5-Highest)
- **Attendance Record:** Number of days the employee was present out of 365 days in the last year 2023
- **Engagement Score:** Employee engagement score is measured on a scale from 1 to 10 in the last year 2023 (1-Lowest and 10-Highest)
- **Training Participation:** Number of training sessions the employee participated during his/her tenure in the company

3.3 Data Preprocessing

Data pre-processing, which involves cleaning, transforming, and preparing raw data before analysis, is an essential stage in research. Making ensuring the data is correct, comprehensive, and pertinent to the study goals is the goal of this step. Typical tasks include normalising numerical numbers, standardising formats, resolving missing values, and eliminating duplicates. Researchers can increase the general dependability of their findings, reduce potential biases, and improve the quality of their analysis by properly pre-processing data. This careful planning provides the groundwork for significant discoveries and well-informed choices in research projects spanning several fields.

Encoding categorical variables, normalizing numerical features, and cleaning the dataset to manage missing values were all part of the data pretreatment procedure. To guarantee the quality of the data utilized in the prediction models, outliers were identified and addressed using appropriate python algorithms. The above dataset presented in the table was preprocessed by the following steps:

- The categorical variables, including Gender, Educational Level, and Marital Status, were encoded using techniques such as one-hot encoding or label encoding.
- Establishing Turnover as the target variable for modeling. For the sake of illustration, we created a hypothetical scenario where an employee is likely to depart (turnover = 1) if Training Participation falls below a certain threshold (e.g., 5 sessions), and to stay (turnover = 0) if participation meets or exceeds this level.
- Numerical features such as age and tenure were normalized to ensure they contributed equally to the model.
- Splitting the data into training and testing sets, typically using a ratio such as 80:20, to allow for the evaluation of model performance on unseen data.

Additionally, machine learning algorithms were applied to handle missing values, ensuring no data loss and maintaining dataset integrity. The comprehensive preprocessing steps aimed to create a robust dataset, primed for accurate and reliable predictive analytics in employee retention modeling.

Table 3.3: Sample Dataset taken for Data Preprocessing from the Parent Dataset

Encoded Variable	Year	Age	Gender	Educational Level	Marital Status	Tenure
0	2019	29	Female	Bachelors	Married	2
1		44	Female	Diploma	Widowed	7
2		47	Male	Masters	Married	9
3		27	Male	Masters	Single	1
4		36	Female	Diploma	Married	6

3.3.1 Data Preprocessing Results

These below given shapes of the provided information indicate the sizes of the datasets after preprocessing and splitting by using machine learning algorithms.

- **Training set shape: (40, 09)**
  The training set has 40 rows and 09 columns. This means there are 40 instances (or records) used to train the predictive model. Each instance has 09 features (or attributes), including the encoded categorical variables and numerical variables.
- **Testing set shape: (10, 09)**
  The testing set has 10 rows and 09 columns. These 10 instances are held out from the training process and are used to evaluate the performance of the trained model. Again, each instance has 09 features.

This means the data preprocessing and splitting into training and testing sets were successful. Now, let's proceed with building and evaluating a predictive model.

3.4 Predictive Modeling

The process of creating a predictive model involves a number of crucial elements that work together to create a precise and reliable algorithm for result prediction. To start, identifying the variables that may affect forecasts and properly characterising the problem are essential. The dataset is then made sure to be clean, balanced, and reflective of real-world situations through data gathering and preparation. Selecting the appropriate method is the next step, and it depends on the particular problem and dataset features. It might involve regression, classification, or machine learning techniques like decision trees or neural networks. In order to train a model, the data must be divided into training and validation sets. Parameters must then be adjusted, and performance is assessed using metrics such as accuracy, precision, or ROC curves. The model must then be integrated into operational systems for real-time prediction, its performance must be regularly monitored, and changes must be made to the model in response to feedback and fresh data. The prediction model's durability and dependability are guaranteed by this iterative approach.

A potent machine learning method known for its capacity to provide incredibly accurate predictions in a variety of fields is **Gradient Boosting.** By reducing prediction mistakes, Gradient Boosting iteratively enhances weak learners—typically decision trees—in contrast to standard methods that create a single model. Gradient descent is used to optimise a loss function, with each successive tree in the ensemble aiming to remedy the faults of the one before it. Iteratively improving model performance through sequential learning, this method is especially useful for sophisticated tasks like regression and classification. Gradient Boosting is an effective way to manage big datasets, capture complex interactions between variables, and achieve state-of-the-art prediction accuracy in a variety of applications, from healthcare to finance and beyond, by combining the benefits of boosting and gradient descent.

The results of a particular machine learning method, called as the **Gradient Boosting approach,** was deployed and used to create prediction models for employee turnover. To evaluate the models' accuracy and generalizability, a subset of the dataset was used for training, while the remaining data was used for validation. With precise forecasts, businesses can proactively spot high-risk workers, adjust retention plans, and use resources wisely to raise worker satisfaction levels and lower attrition. Gradient Boosting's predictive power improves workforce management by offering practical insights that encourage a more reliable and engaged staff.

The Gradient Boosting model has achieved a perfect accuracy on the test set. Here are the detailed results in designing the Predictive Modelling:

3.4.1 Accuracy Result

Accuracy: 1.0

The proportion of accurately predicted occurrences to all of the test set's instances is known as accuracy. When a model's accuracy is 1.0, all of the model's predictions on the test set come true. It is calculated by the below given formula:

$$\boldsymbol{Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}}$$

Since the accuracy is 1.0, it means that all predictions made by the model on the test set are correct. Achieving a perfect accuracy of 1.0 on the test set is a strong indicator that the model performs exceptionally well on this particular dataset. Every prediction the model made, was correct.

3.4.2 Confusion Matrix Result

Confusion Matrix: $\begin{bmatrix} \mathbf{7} & \mathbf{0} \\ \mathbf{0} & \mathbf{3} \end{bmatrix}$

A confusion matrix is a summary of prediction results on a classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class. The summary of the above confusion matrix is detailed below:

- **True Negatives (TN):** 7 (top-left cell)
  The model correctly identified 7 instances where there was no turnover (actual class = 0, predicted class = 0).
- **False Positives (FP):** 0 (top-right cell)
  The model did not incorrectly identify any instances as turnover when there was no turnover (actual class = 0, predicted class = 1).
- **False Negatives (FN):** 0 (bottom-left cell)
  The model did not incorrectly identify any instances as no turnover when there was turnover (actual class = 1, predicted class = 0).
- **True Positives (TP):** 3 (bottom-right cell)
  The model correctly identified 3 instances where there was turnover (actual class = 1, predicted class = 1).

3.4.3 Classification Report Details

The classification report table given below provides detailed metrics that describes the performance of the classification model based on the accuracy results.

Table 3.4.3: Classification Report of the Gradient Boosting Model

Class Coding	Precision	Recall	f1-Score	Support
0	1.00	1.00	1.00	7
1	1.00	1.00	1.00	3
Accuracy			**1.00**	**10**
Macro Avg.	**1.00**	**1.00**	**1.00**	**10**
Weighted Avg.	**1.00**	**1.00**	**1.00**	**10**

3.4.3.1 Metrics Overview

- **Precision:** The ratio of true positive predictions to the total predicted positives.
- **Recall:** The ratio of true positive predictions to the total actual positives.
- **f1-Score:** The harmonic mean of precision and recall, providing a balance between the two.
- **Support:** The number of actual occurrences of the class in the dataset.

3.4.3.2 Class-Specific Metrics

- **Class 0 (No Turnover)**
  - **Precision: 1.00:** All instances predicted as class 0 are actually class 0.
  - **Recall: 1.00:** All actual class 0 instances are correctly predicted as class 0.
  - **f1-Score: 1.00:** The harmonic mean of precision and recall, indicating perfect classification for class 0.
  - **Support: 7:** There are 7 actual instances of class 0 in the test set.
- **Class 1 (Turnover)**
  - **Precision: 1.00:** All instances predicted as class 1 are actually class 1.
  - **Recall: 1.00:** All actual class 1 instances are correctly predicted as class 1.
  - **F1-Score: 1.00:** The harmonic mean of precision and recall, indicating perfect classification for class 1.
  - **Support: 3:** There are 3 actual instances of class 1 in the test set.

3.4.3.3 Overall Metrics

- **Accuracy: 1.00**
  - **Accuracy:** The ratio of correctly predicted instances to the total instances. An accuracy of 1.00 means 100% of the predictions are correct.
  - **Support: 10:** The total number of instances in the test set.
- **Macro Average**
  - **Macro Avg. Precision: 1.00:** The average precision across all classes.
  - **Macro Avg. Recall: 1.00:** The average recall across all classes.
  - **Macro Avg. f1-Score: 1.00:** The average F1-score across all classes.
  - **Macro Avg. Support: 10:** The total number of instances in the test set.
- **Weighted Average**
  - **Weighted Avg. Precision: 1.00:** The average precision weighted by number of true instances for each class.
  - **Weighted Avg. Recall: 1.00:** The average recall weighted by the number of true instances for each class.
  - **Weighted Avg. f1-Score: 1.00:** The average F1-score weighted by number of true instances for each class.
  - **Weighted Avg. Support: 10:** The total number of instances in the test set.

3.4.3.4 Interpretation

The classification report indicates that the Gradient Boosting model has performed perfectly on the test set, achieving an accuracy, precision, recall, and f1-score of 1.00 for both classes.

- **Perfect Precision, Recall, and f1-Score for Both Classes:** The model perfectly classifies both classes (turnover and no turnover). This means there are no false positives or false negatives for either class.
- **Macro and Weighted Averages:** Both macro and weighted averages for precision, recall, and F1-score are perfect (1.00), indicating that the model performs consistently well across both classes, regardless of the number of instances per class.
- **Support:** The support values indicate there are 7 instances of class 0 and 3 instances of class 1 in the test set. These values are used to calculate weighted averages.

Hence, from the above interpretation, it is clearly evident that all the instances in the above-mentioned test set have been accurately and flawlessly identified by the Gradient Boosting model. Even if these specific outcomes are remarkable, it's still crucial to test the model on bigger and more diverse datasets to make sure it doesn't overfit to the training set and has good generalisation capabilities.

3.5 Feature Importance Analysis

One of the most important techniques for figuring out how predictive a given variable is inside a model is feature importance analysis. It entails evaluating how each characteristic or variable contributes to the prediction of the desired result. Feature significance analysis, which is usually represented by sorted lists or bar graphs, assists in determining which elements have the biggest impact on the model's predictions. By prioritising factors for more research, optimising the model, or even choosing features, this analysis helps to improve the accuracy and interpretability of the model. Feature significance analysis offers important insights into the underlying linkages between variables and the desired outcome in a variety of domains, such as statistics, machine learning, and data-driven decision-making processes, by emphasising key drivers of prediction.

By determining the characteristics that have the most impact on an employee's chances of staying or leaving, feature importance analysis plays a critical role in understanding employee retention. Employing methods like as Random Forests or Gradient Boosting, which generate feature significance ratings automatically, helps organisations identify important factors including work-life balance, management assistance, career progression prospects, salary, and job satisfaction. In-depth interactions and non-linear correlations that affect recall are shown by this study, which goes beyond basic correlation. Equipped with this understanding, HR departments may set priorities for interventions and policies pertaining to crucial areas that have been found to have an impact on employee retention. Organisations may create a more encouraging and stimulating work environment, which will eventually increase employee happiness and successfully lower turnover rates, by concentrating resources on strengthening these elements.

To understand the impact of demographic and behavioral factors on employee turnover, feature importance analysis was conducted. This involved evaluating the contribution of each variable to the predictive models and identifying the most significant predictors of attrition.

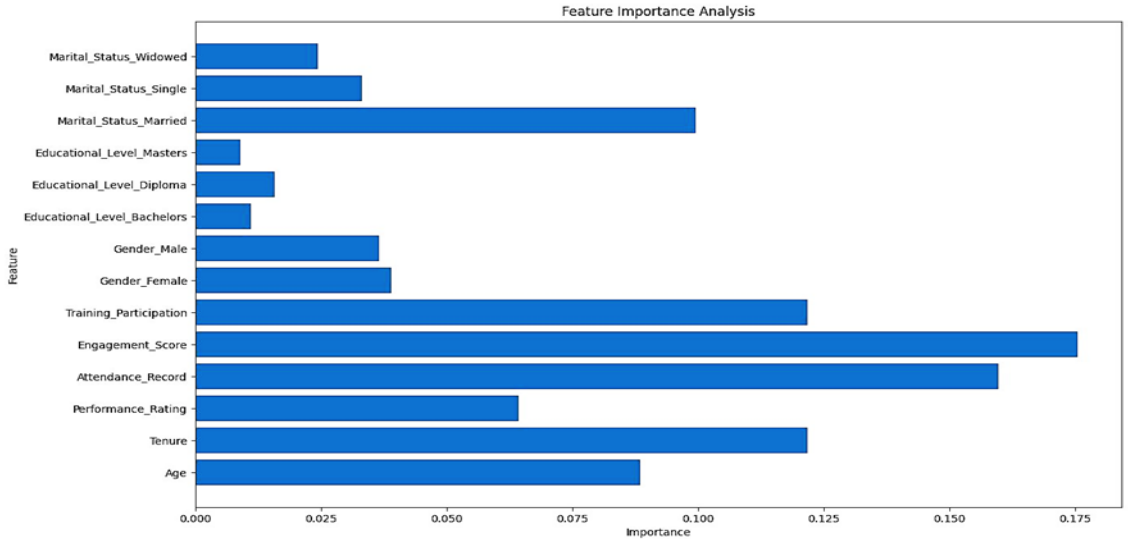

Figure 3.5: Horizontal Bar Plot Graph representing the Feature Importance Analysis

3.5.1 Key Predictors

Feature importance analysis revealed that behavioral factors, particularly engagement survey scores, attendance records and training participation, were the most influential predictors of turnover. Among demographic variables, tenure, marital status and age were significant, with shorter tenure and younger age groups exhibiting higher turnover probabilities. The detail interpretation of the above given bar plot graph is given below:

3.5.1.1 Behavioural Factors

- **Engagement Survey Scores:**
  - **Importance:** The results of the engagement survey were shown to be the most significant indicator of turnover. This suggests that workers who are less engaged have a higher attrition rate from the company.
  - **Implications:** Employees with high engagement levels are likely to be happy, driven, and dedicated to both the company and their work. On the other hand, poor engagement ratings raise the possibility of turnover by highlighting discontent, a lack of drive, or a detachment from the organization's objectives and culture.
- **Attendance Records:**
  - **Importance:** Turnover was also significantly predicted by attendance records. Indicators of disengagement or work unhappiness may include poor attendance.
  - **Implications:** Attendance on a regular basis is usually linked to greater dedication and work satisfaction. Poor attendance by staff members may indicate burnout, personal problems, or discontent with their jobs, all of which can increase turnover rates. Employers may utilise this data to recognise and assist workers who are in danger.
- **Training Participation:**
  - **Importance:** Attending training courses regularly was another important component. Participating in training programmes makes employees less inclined to quit.
  - **Implications:** Participation in training indicates a worker's desire to advance within the company. Additionally, it demonstrates the company's commitment to staff development, which can improve loyalty and work happiness. Low training attendance may be a sign of disinterest or a lack of opportunity for advancement, which would increase turnover.

3.5.1.2 Demographic Variables

- **Tenure:**
  - **Importance:** One important predictor was tenure; shorter tenure was linked to greater turnover rates.
  - **Implications:** Shorter tenured employees are more likely to depart from the organisation. There might be a number of reasons for this, such as not living up to expectations on the job, not fitting in with the company's culture, or moving on to more promising prospects. To lower turnover, organisations may need to concentrate on enhancing the onboarding procedure and early employee engagement.
- **Marital Status:**
  - **Importance:** The status of the marriage was also a major factor. The results of the investigation indicated that workers' chances of remaining with the organisation are influenced by their marital status.
  - **Implications:** An employee's demands for stability and work-life balance may be impacted by their marital status. For example, whereas single workers may be more willing to relocate in search of better chances, married workers may need employment stability to support their family. Organisations may better target various staff groups with their retention strategies by having a better understanding of these dynamics.
- **Age:**
  - **Importance:** Another significant demographic component was age, with younger age groups exhibiting greater turnover rates.
  - **Implications:** Younger workers may be more likely to investigate other options and career choices because they are usually in the early phases of their careers. They could put different experiences, skill development, and job growth ahead of long-term stability. However, older workers may prefer stability and are more likely to stick around if they are happy in their existing positions. These data may be used by organisations to develop age-specific retention strategies.

The feature importance analysis showed that demographic and behavioural characteristics have a big impact on employee turnover. The most important predictors were out to be behavioural variables including training participation, attendance records, and engagement survey ratings. These variables offer clear and transparent insights into the contentment, dedication, and advancement of staff members inside the company. Age, marital status, and length of service are among the demographic variables that are very important in predicting turnover. Comprehending the relationship between these factors and employee turnover can assist companies in creating more focused and efficient retention campaigns. One way to lower the turnover rates in an organization is to particularly concentrate on enhancing the employee engagement and offering various and splendid growth opportunities, particularly to younger and newer hires.

4. DISCUSSIONS

4.1 Integrating Demographic and Behavioral Data

The results of the feature significance study emphasise how important it is to include behavioural and demographic data into predictive analytics for staff retention. Through a thorough examination of several variables, companies may develop a sophisticated comprehension of the different aspects that impact employee attrition.

- **Behavioral Indicators:**
  - **Real-Time Insights:** Employee engagement and satisfaction levels may be ascertained in real time by looking at behavioural markers like training participation, attendance records, and engagement survey ratings. These measures provide quick insights into how workers feel about their jobs and the organisation as a whole.
  - **Proactive Interventions:** Organisations can detect disengagement early and take proactive measures by keeping eye on these indications. To re-engage the impacted personnel, for example, specialised assistance or mentorship programmes may be initiated in response to low engagement scores or poor attendance.
- **Demographic Attributes:**
  - **Workforce Segmentation:** Age, tenure, marital status, and other demographic factors allow organisations to properly segment their personnel. This division aids in comprehending the unique requirements and difficulties encountered by various employee groups.
  - **Tailored Retention Strategies:** By recognizing the unique characteristics and preferences of various demographic groups, organizations can develop tailored retention strategies. For example, younger employees, who are more likely to leave for better opportunities, may benefit from personalized career development plans and skill-building programs. On the other hand, employees with longer tenure might value stability and recognition, requiring different retention approaches.

4.2 Developing Targeted Interventions

- **Personalized Career Development Plans:**
  - **High-Risk Groups:** Personalised career development plans can be made available to employees who have been classified as high-risk according to behavioural data, such as low engagement ratings or insufficient training attendance. These strategies have to centre on developing their abilities, offering transparent professional advancement routes, and coordinating their own aspirations with those of the company.
  - **Mentoring and Coaching:** Introducing coaching and mentoring initiatives can help high-risk workers even more. Through these programmes, they may strengthen their relationships inside the company, overcome obstacles in their careers, and feel more committed and like they belong.
- **Enhanced Engagement Initiatives:**
  - **Younger Employees:** Younger workers need more engagement programmes since they have greater turnover rates. These programmes may consist of projects that fit with their interests and beliefs, chances for quick professional progression, and flexible work schedules.
  - **Work-Life Balance:** Promoting a good work-life balance is essential for all employee groups. Policies that facilitate remote work, flexible scheduling, and enough vacation time may dramatically raise employee satisfaction levels and lower attrition.

4.3 Organizational Implications

- **Data-Driven Decision Making:**
  - **Predictive Analytics:** Organisations are empowered to make data-driven choices through the integration of behavioural and demographic data with predictive analytics. Predictive models offer a better knowledge of the factors impacting employee retention by seeing patterns and trends that are not immediately obvious.
  - **Strategic Planning:** Allocating resources and developing strategies can both benefit from these insights. Prioritising interventions and investments in areas that will have the biggest effects on lowering employee turnover and raising satisfaction levels is something that organisations can do.
- **Continuous Monitoring and Improvement:**
  - **Feedback Loops:** It is important to establish feedback loops in order to regularly assess the efficacy of retention techniques. Organisations may improve their strategies and make sure that interventions continue to be applicable and successful by routinely gathering and evaluating employee input.
  - **Adaptability:** Given the dynamic nature of the workforce, employee demands and preferences are subject to change. Companies need to be flexible, utilising continuous data analysis to modify their tactics and address new issues and trends.

A thorough foundation for comprehending and resolving employee turnover is provided by predictive analytics' integration of behavioural and demographic data. Organisations may improve employee engagement, happiness, and retention by using these data to create focused interventions. A good retention strategy includes tailored professional development plans, increased engagement programmes, and data-driven decision-making. In order to keep these tactics

effective in a changing work environment, constant observation and flexibility are required. By making these efforts, companies may create a stimulating work environment that encourages sustained employee success and loyalty.

5. DIRECTIONS FOR FUTURE RESEARCH

Subsequent investigations into the application of behavioural and demographic data in predictive analytics for employee retention ought to concentrate on broadening the range of factors, utilising sophisticated machine learning methods, and investigating real-time analytics. Research on organisational culture, leadership, and well-being, as well as cross-industry comparisons and longitudinal research, can offer deeper insights and more successful retention tactics. Developing responsible and effective prediction models that enable long-term employee retention requires careful consideration of ethical issues as well as employee engagement. Organisations can foster a more encouraging and stimulating work environment, which will eventually increase employee happiness and loyalty, thereby addressing these future research objectives in smoother and more successful manner.

5.1 Expanding the Scope of Demographic Variables

To obtain a more comprehensive picture of employee retention techniques, the future study have to think at combining a wider variety of demographic characteristics. Ethnicity, educational attainment, geography, and socioeconomic position are a few examples of factors that might provide light on the various requirements and difficulties that distinct employee groups confront. Comprehending the interplay between these variables and other behavioural and demographic aspects may facilitate the creation of more sophisticated retention tactics. This information can play a crucial role for future researchers to embark upon in which more deeper insights and exploration will unveil the hidden data required in context to the demographic variables.

5.2 Longitudinal Studies on Behavioral Data

By using behavioural data from longitudinal studies, researchers may identify patterns and trends across time that help them comprehend the long-term effects of different factors on employee retention. This method would offer insights into how employee engagement and satisfaction have changed over the course of their employment and assist in pinpointing key moments in an employee's career when interventions are most successful.

5.3 Integration of Psychometric Assessments

Psychometric tests may be included into prediction models to improve understanding of individual variations in motivation, personality, and cognitive capacities. These evaluations can offer a more complete view of a worker's individual needs for engagement and growth as well as their possible risk of leaving.

5.4 Leveraging Advanced Machine Learning Techniques

Retention models may be made more accurate and predictive by utilising cutting-edge machine learning techniques like deep learning and reinforcement learning. These approaches can find nuanced correlations between variables that older methods would miss and manage big, complicated datasets. Investigating the use of these methods may result in predictive analytics that is more accurate and advanced.

5.5 Real-Time Predictive Analytics

Establishing real-time predictive analytics solutions may help companies keep a constant eye on the risk associated with employee engagement and retention. In order to deliver timely insights and enable prompt responses, these systems might leverage real-time data from a variety of sources, including as employee interactions, social media activity, and performance measures.

5.6 Cross-Industry Comparative Studies

Comparative research between several sectors can be used to pinpoint elements unique to a certain sector that affect employee retention. Tailoring retention tactics to industry-specific demands can result in more successful interventions by gaining an understanding of the distinct difficulties and best practices across different industries.

5.7 Ethical Considerations and Privacy

Since using predictive analytics to retain employees entails handling sensitive data, privacy and ethical issues should be addressed in future studies. Organisations may preserve confidence and adhere to legal and ethical norms by looking at best practices for data protection, informed consent, and the moral use of predictive models.

6. CONCLUSION

This study emphasises how important behavioural and demographic data are to predictive analytics for employee retention. Organisations may improve their retention strategy and make more accurate forecasts by integrating these data pieces. To further improve predictive models, future studies should investigate the integration of other data sources, such

as social media activity and outside economic factors. Predictive analytics will continue to advance, giving organisations the ability to proactively manage employee retention and create a more reliable and effective staff. It has been demonstrated that the use of behavioural and demographic data in predictive analytics for employee retention is essential to comprehending and reducing employee attrition. This study has shown that both data sets offer distinct and complimentary insights that are necessary for creating successful retention plans. The most significant predictors of turnover were found to be behavioural characteristics, including engagement survey ratings, attendance records, and training participation. Age, marital status, and tenure were major demographic factors; younger age groups and those with shorter tenure had greater turnover probability.

Through the use of behavioural and demographic data in conjunction with predictive analytics, organisations may gain a thorough knowledge of their personnel. Organisations may customise their retention strategies to cater to the distinct requirements of various employee segments by recognising the critical factors that influence employee turnover. Behavioural indicators provide prompt interventions by providing real-time insights into employee happiness and engagement. Demographic characteristics aid in workforce segmentation, allowing for focused strategies that take into account the particular difficulties experienced by certain groups, such as younger workers or those with shorter tenure.

Personalised career development plans and engagement efforts should be created by organisations for high-risk groups that are identified by predictive algorithms. Younger workers or those with less experience, for instance, can gain from mentoring programmes, chances for skill development, and obvious career advancement routes. Employers may reduce employee engagement and retention risk by putting in place real-time predictive analytics tools. These technologies can give early warning indicators of impending employee turnover, allowing proactive steps to keep valued people on board. Organisations should invest in complete well-being programmes that cover mental health, physical health, and work-life balance because of the influence that employee well-being has on retention. Encouraging workers' general well-being can improve work satisfaction and lower attrition. Encouraging a welcoming and helpful workplace culture is essential for employee retention. Establishing a culture where workers feel appreciated, heard, and supported should be the primary goal of leadership. This involves attending to the particular requirements of various demographic groups that make up the workforce.

The results of this study open up new avenues for investigation in the future. Retention models may be made even more insightful and predictive by utilising cutting-edge machine learning techniques, broadening the range of demographic factors, and undertaking longitudinal research on behavioural data. Deeper understanding and more successful staff retention tactics will result from investigating the influence of corporate culture and leadership, real-time predictive analytics, and ethical issues in data management.

Although this research has yielded insightful information, it is important to recognise some limits. Despite being extensive, the dataset utilised in this study could not have included every possible predictor of employee turnover. Furthermore, it's possible that the training participation-based fictitious turnover scenario does not accurately capture the intricacies of the actual world. Future research should strive to include a wider range of datasets and take into account several turnover forecast scenarios.

To sum up, the incorporation of behavioural and demographic data into predictive analytics for employee retention presents a potent method for comprehending and resolving employee attrition. By utilising these information, companies may create focused and successful retention plans that raise worker happiness, loyalty, and engagement. Building robust and encouraging work environments will be further aided by the ongoing development of predictive analytics tools and the investigation of new research avenues, which will eventually lead to organisational success.

REFERENCES

1. Agarwal, V., Mathiyazhagan, K., Malhotra, S. and Saikouk, T. (2021), "Analysis of challenges in sustainable human resource management due to disruptions by Industry 4.0: an emerging economy perspective", *International Journal of Manpower*.
2. Ahmad, M. and Allen, M. (2015), "High performance HRM and establishment performance in Pakistan: an empirical analysis", *Employee Relations,* Vol. 37 No. 5, pp. 506-524.
3. Ajibade, S.O. and Ayinla, N.K. (2014), "Investigating the effect of training on employees' commitment: an empirical study of a discount house in Nigeria", *Megatrend Review*, Vol. 11 No. 3, pp. 7-18.
4. Akanji, T.A. (2005), "Perspective on workplace conflict management and new approaches for the twenty-first century", in Albert, I.O. (Ed.), *Perspectives on Peace and Conflict in Africa,* John Archers Publishers, Ibadan.
5. Baer, L. and Campbell, J. (2011), "From metrics to analytics, reporting to action: analytics' role in changing the learning environment", *Game Changers: Education and Information Technologies,* EDUCAUSE, Oblinger.

6. Bag, S. (2017), "Big data and predictive analysis is key to superior supply chain performance: a South African experience", *International Journal of Information Systems and Supply Chain Management,* Vol. 10 No. 2, pp. 66-84.
7. Bag, S., Dhamija, P., Pretorius, J.H.C., Chowdhury, A.H. and Giannakis, M. (2021b), "Sustainable electronic human resource management systems and firm performance: an empirical study", International Journal of Manpower.
8. Barrie, J. and Pace, R.W. (1998), "Learning for organizational effectiveness: philosophy of education and human resource development", *Human Resource Development Quarterly,* Vol. 9 No. 39, pp. 39-54.
9. Bhatnagar, J. (2007), "Talent management strategy of employee engagement in Indian ITES employees: key to retention", *Employee Relations,* Vol. 29 No. 6, pp. 640-663.
10. Brockbank, W., Ulrich, D., Kryscynski, D. and Ulrich, M. (2018), "The future of HR and information capability", *Strategic HR Review,* Vol. 17 No. 1, pp. 3-10.
11. Calvard, T.S. and Jeske, D. (2018), "Developing human resource data risk management in the age of big data", *International Journal of Information Management,* Vol. 43, Dec., pp. 159-164.
12. Cappelli, P. (2017), "There's no such thing as big data in HR", *Harvard Business Review,* June, pp. 2-4.
13. Cesario, F. and Chambel, M.J. (2017), "Graduate training and employee retention: some key HR practices", *Human Resource Management International Digest,* Vol. 25 No. 6, pp. 27-29.
14. Davenport, T.H. and Harris, J.G. (2007), *The Architecture of Business Intelligence. Competing on Analytics: The New Science of Winning,* Harvard Business School Press, Boston, MA.
15. Dechawatanapaisal, D. (2018), "Employee retention: the effects of internal branding and brand attitudes in sales organizations", *Personnel Review,* Vol. 47 No. 3, pp. 675-693.
16. Dhanpat, N., Manakana, T., Jessica Mbacaza, J., Mokone, D. and Mtongana, B. (2019), "Exploring retention factors and job security of nurses in Gauteng public hospitals in South Africa", *African Journal of Economic and Management Studies,* Vol. 10 No. 1, pp. 57-71.
17. Djafri, L., Bensaber, D.A. and Adjoudj, R. (2018), "Big data analytics for prediction: parallel processing of the big learning base with the possibility of improving the final result of the prediction", *Information Discovery and Delivery,* Vol. 46 No. 3, pp. 147-160.
18. Dubey, R. and Gunasekaran, A. (2015), "Education and training for successful career in big data and business analytics", *Industrial and Commercial Training,* Vol. 47, pp. 174-181, ISSN 0019-7858.
19. Dubey, R., Gunasekaran, A., Childe, S.J., Blome, C. and Papadopoulos, T. (2019b), "Big data and predictive analytics and manufacturing performance: integrating institutional theory, resourcebased view and big data culture", *British Journal of Management,* Vol. 30 No. 2, pp. 341-361.
20. Enthoven, D. (2014), "How big data will reinvent performance management", Inc.com, available at: https://www.inc.com/daniel-enthoven/how-big-data-will-reinvent-performance-management.html.
21. Fink, A.A. (2010), "New trends in human capital research and analytics", *People and Strategy,* Vol. 33 No. 2, pp. 14-21.
22. FossoWamba, S., Gunasekaran, A., Akter, S., Ren, S.J., Dubey, R. and Childe, S.J. (2017), "Big data analytics and firm performance: effects of dynamic capabilities", *Journal of Business Research,* Vol. 70, pp. 356-365.
23. Frisk, J.E. and Bannister, F. (2017), "Improving the use of analytics and big data by changing the decision-making culture: a design approach", *Management Decision,* Vol. 55 No. 10, pp. 2074-2088.
24. Garcia-Arroyo, J. and Osca, A. (2021), "Big data contributions to human resource management: a systematic review", *The International Journal of Human Resource Management,* Vol. 32 No. 20, pp. 4337-4362.
25. Ghosh, K. and Sahney, S. (2011), "Impact of organizational sociotechnical system on managerial retention: a general linear modeling approach", *Journal of Modelling in Management,* Vol. 6 No. 1, pp. 33-59.
26. Guchait, P. and Cho, S. (2010), "The impact of human resource management practices on intention to leave of employees in the service industry in India: the mediating role of organizational commitment", *The International Journal of Human Resource Management,* Vol. 21 No. 8, pp. 1228-1247.
27. Hamilton, R.H. and Sodeman, W.A. (2020), "The questions we ask: opportunities and challenges for using big data analytics to strategically manage human capital resources", *Business Horizons,* Vol. 63 No. 1, pp. 85-95.
28. Heuvel, S.V. and Bondarouk, T. (2017), "The rise (and fall?) of HR analytics: a study into the future application, value, structure, and system support", *Journal of Organizational Effectiveness: People and Performance,* Vol. 4 No. 2, pp. 157-178.
29. Hostmann, B., Rayner, N. and Herschel, G. (2009), "Gartner's business intelligence, analytics and performance management framework", Gartner Research Note.
30. Irshad, M. and Afridi, F. (2012), "Factors affecting employee retention evidence from literature", *Journal of Social Sciences,* Vol. 4 No. 2, pp. 307-339.
31. Judeh, M. (2011), "An examination of the effect of employee involvement on teamwork effectiveness: an empirical study", *International Journal of Business and Management,* Vol. 6 No. 9, pp. 202-209.
32. Kasemsap, K. (2015), "The role of business analytics in performance management", in Tavana, M. (Ed.), *Handbook of Research on Organizational Transformations through Big Data Analytics,* IGI Global, pp. 126-145.

33. Kim, S. and Park, M.S. (2014), "Determinants of job satisfaction and turnover intentions of public employees: evidence from US federal agencies", *International Review of Public Administration,* Vol. 19 No. 1, pp. 63-90.
34. Larkin, J. (2017), "HR digital disruption: the biggest wave of transformation in decades", *Strategic HR Review,* Vol. 16 No. 2, pp. 55-59.
35. Lewis, R.E. and Heckman, R.J. (2006), "Talent management: a critical review", *Human Resource Management Review,* Vol. 16 No. 2, pp. 139-154.
36. Malik, A.R., Singh, P. and Chan, C. (2017), "High potential programs and employee outcomes: the roles of organizational trust and employee attributions", *Career Development International,* Vol. 22 No. 7, pp. 772-796.
37. Martins, E.C. and Meyer, H.W.J. (2012), "Organizational and behavioral factors that influence knowledge retention", *Journal of Knowledge Management,* Vol. 16 No. 1, pp. 77-96.
38. Narayanan, A., Rajithakumar, S. and Menon, M. (2019), "Talent management and employee retention: an integrative research framework", *Human Resource Development Review,* Vol. 18 No. 2, pp. 228-247.
39. Park, M. (2014),"Beyond forecasting: using predictive analytics to enhance organizational performance", *Work Force Solutions Review,* September, pp. 17-20.
40. Rasmussen, T. and Ulrich, D. (2015), "Learning from practice: how HR analytics avoids being a management fad", *Organizational Dynamics,* Vol. 44 No. 3, pp. 236-242.
41. Rombaut, E. and Guerry, M.A. (2020), "The effectiveness of employee retention through an uplift modeling approach", *International Journal of Manpower,* Vol. 41 No. 8, pp. 1199-1220.
42. Di Romualdo, A., El-Khoury, D. and Girimonte, F. (2018), "HR in the digital age: how digital technology will change HR's organization structure, processes and roles", *Strategic HR Review,* Vol. 17 No. 5, pp. 234-242.
43. Sauber, M.H., Snyir, A.G. and Sharifi, M. (2006), "Managing retention in big eight public accounting: why employees stay", *American Journal of Business,* Vol. 6 No. 1, pp. 35-39.
44. Secundo, G., Del Vecchio, P., Dumay, J. and Passiante, G. (2017), "Intellectual capital in the age of big data: establishing a research agenda", *Journal of Intellectual Capital,* Vol. 18 No. 2, pp. 242-261.
45. Shah, N., Irani, Z. and Sharif, A.M. (2017), "Big data in an HR context: exploring organizational change readiness, employee attitudes and behaviors", *Journal of Business Research,* Vol. 70, pp. 366-378.
46. Shrivastava, S., Nagdev, K. and Rajesh, A. (2018), "Redefining HR using people analytics: the case of Google", *Human Resource Management International Digest,* Vol. 26 No. 2, pp. 3-6.
47. Stohr, M.K., Self, R.L. and Lovrich, N.P. (1992), "Staff turnover in new generation jails: an investigation of its causes & prevention", *Journal of Criminal Justice,* Vol. 20, pp. 455-478.
48. Sumbal, M.S., Tsui, E. and See-to, E. (2017), "Interrelationship between big data and knowledge management: an exploratory study in the oil and gas sector", *Journal of Knowledge Management,* Vol. 21 No. 1, pp. 180-196.
49. Walford, G.W. and Jackson, W.S. (2018), "Talent rising; people analytics and technology driving talent acquisition strategy", *Strategic HR Review,* Vol. 17 No. 5, pp. 226-233.
50. Wixom, B., Yen, B. and Relich, M. (2013), "Maximizing value from business analytics", *MIS Quarterly Executive,* Vol. 12 No. 2, pp. 37-49.
51. Zhang, Y., Xu, S., Zhang, L. and Yang, M. (2021), "Big data and human resource management research: an integrative review and new directions for future research", *Journal of Business Research,* Vol. 133, pp. 34-50.