

# AttriSense: Data-Driven Analysis for Employee Attrition

Arpana Sinha<sup>1</sup>

*Department of Computer Science and Engineering, Sikkim Manipal Institute of Technology, SMU, Majitar, Sikkim, India.*

**Abstract**—Organizations have a major problem from employee attrition, which affects productivity and raises recruitment expenses. The objective of this project is to forecast employee attrition through the use of multiple machine learning models, such as Artificial Neural Networks, Random Forest and Support Vector Machines. The models were trained by examining employee-related characteristics from the IBM HR Analytics dataset, including job role, satisfaction level and performance rating in order to detect trends that result in attrition. HR departments may lower turnover and retain important personnel by using this predictive technique to drive their decisions.

**Index Terms**—Employee Attrition, IBM HR Analytics dataset, Machine Learning Models- Linear Regression, Random Forest, K-Nearest Neighbor, ANN, RNN, Data Science, Predictive Modeling.

## I. INTRODUCTION

The progressive loss of workers over time or staff attrition is a recurring problem for contemporary businesses. High turnover rates impair team performance and disturb operational continuity in addition to raising the expense of hiring and training new employees. HR departments may improve workforce stability and take preventative action by knowing and anticipating the variables that contribute to employee loss.



Fig: Employee Attrition Reasons

Now that the fields of machine learning (ML) and data analytics have progressed, it is possible to develop predictive algorithms that examine employee data in order to estimate the probability of attrition. Through the use of several supervised machine learning algorithms, this research seeks to forecast employee attrition based on a variety of factors, including job rating, years at the organization, work satisfaction, overtime and salary.

My methodology involves preprocessing the IBM HR Analytics Employee Attrition dataset and applying models like Logistic Regression, Random Forest, Decision Tree, K-Nearest Neighbours (KNN), Artificial Neural Network (ANN) and Recurrent Neural Network (RNN). Evaluation criteria including as accuracy, precision, recall and F1-score are used to assess the performance of different models in order to identify the best predictor.

In addition to demonstrating the usefulness of machine learning in HR analytics, this study attempts to offer practical insights that can assist firms in retaining talent and raising employee engagement. With its emphasis on the part that tenure, job role & overtime play in employee attrition, the ANN model proved to have been the most accurate and balanced predictor. These findings can help HR professionals create strategic actions to improve organizational efficiency and reduce attrition.

## II. LITERATURE SURVEY

Employee attrition reasons and the efficacy of prediction strategies in human resource management have been examined in a number of research and analytical models. A carefully chosen overview of pertinent research that focuses on data-driven methods for comprehending and reducing employee turnover is provided below, highlighting the field's changing landscape.

*A. Shahin Manafi & Francesco Pattarin et al(2025)- Predicting Employee Attrition and explaining its determinants*

This study uses actual data from a significant Italian financial organization to examine staff attrition. To figure out which employees are likely to quit, it relies on machine learning algorithms, particularly Random Forest. The study pinpoints the main variables that affect attrition, including pay, age, job function, tenure and working hours. It underlines the necessity of customized HR strategies and illustrates the ways in which organizational, job-related and personal factors affect turnover. According to the results, data-driven analysis can help with more individualized and successful retention tactics, particularly in the changing post-pandemic workplace.

*B. M. Lakshmanam & R. Rogini Kala et al(2024)- Impact of Employee Attrition rate in industry*

This research examines how employee attrition affects business performance, emphasizing problems such as decreased morale, higher hiring expenses, knowledge loss and lost productivity. A small amount of attrition fosters creativity and new talent, but a large amount of turnover harms team chemistry and long-term objectives. Key factors that contribute to attrition are identified by the study, including low pay, limited career advancement, and inefficient leadership. By using data analysis, it also demonstrates a strong correlation between assigning work and employee happiness. Better data analysis, enhanced feedback systems and tailored retention initiatives are some of the tactics that the study recommends to reduce turnover and promote a steady, effective workplace. Systems and tailored retention initiatives are some of the tactics the study recommends to reduce turnover and promote a steady, effective workplace.

*C. Dr. Sankararaman, Dr. Ragothaman, Balakrishnan, Divya Barathy, Dineshsurya et al (2023) – A study on Understanding Employee Attrition- Causes and Consequences in the workplace.*

From this study I found that the main causes of employee turnover and the effects it has on the firm, as determined by a poll of 110 workers. It names high levels of job stress, a lack of opportunity for advancement and a poor work-life balance as the main causes of attrition. According to the study, job happiness is negatively connected with more overtime and stress, but work-life balance and total employee

satisfaction are positively impacted by good organizational support for stress management. It was discovered that demographic variables including age, gender and marital status had no discernible impact on total job satisfaction, indicating that the underlying reasons for turnover are more widespread. The report suggests that in order to retain staff and boost morale and productivity at work, HR managers should put in place policies that promote work-life balance, offer possibilities for career advancement, and implement stress-reduction initiatives.

*D. Kishori Singh, Reetu Singh et al(2019) – A Study on Employee Attrition: Effects & Causes*

This study investigates attrition's complex origins and effects on businesses and workers. It looks at things like economic conditions, employment stagnation and business atmosphere from an internal, external, as well as individual perspective. According to the study, there are drawbacks, such as decreased morale and productivity, as well as possible advantages, such as cost effectiveness and new talent. It highlights the importance of emotional intelligence and leadership styles in attrition management and provides solutions to improve retention and workplace satisfaction, such as skill development, employee engagement and flexible policies.

### III. FACTORS INFLUENCING EMPLOYEE ATTRITION

*A. Workplace Contentment*

- *Lack of Job Fulfillment:* If workers fail to find their jobs fulfilling, they might quit, which ultimately results in attrition and discontent.
- *Work which is tedious or boring:* Workers may look for more interesting positions if they get offered repetitive tasks with little variety or potential for advancement.

In short Employees that are dissatisfied with their jobs may seek out more interesting opportunities.

*B. Chances for Professional Growth*

- *Limited Opportunities for Growth:* Employees can leave for companies that provide development opportunities if they believe that they have very little chance of professional growth or advancement.

- *Lack of Training or Skill Development:* Employees may feel stagnate and decide to leave if they find there aren't enough opportunities for learning and skill development.

In short Employees who have few opportunities for development and promotion seek companies that provide skill development and career progression.

#### C. Compensation & Benefits

- *Low Salary:* Workers could be looking for better-paying substitutes elsewhere if they consider that their salary are not being fairly compensated as per their work.
- *Inadequate Benefits:* Employees who are unhappy with their health insurance, retirement programs, or other perks may seek for employers who provide more competitive or appealing benefits.

In short low salary or inadequate benefits (health, retirement plans) can lead employees to leave for better compensation.

#### D. Job Security

- *Insecure employment:* Employees are more inclined to seek more stable job opportunities if they feel their position is insecure as a result of firm unpredictability, frequent job cuts, or downsizing.
- *Absence of a Clear Career Path:* Workers may look for more secure positions elsewhere if they do not perceive a clear, steady career path or the path of their interest.

In short uncertainty brought on by frequent layoffs or unstable businesses makes workers more likely to look for more secure work.

#### E. Organization Culture

- *Toxic Workplace:* Employees may decide to quit an organization with a toxic or bad work culture that is marked by high conflict, a lack of trust, or poor relationships among coworkers.
- *Absence of Diversity or Inclusion:* Workers may decide to leave for more inclusive workplaces if they do not feel valued or included due to their identity, background, or views.

In short employees may feel underappreciated and attrition may increase in a poisonous or intolerant company culture.

## IV. DIFFERENT KINDS OF EMPLOYEE ATTRITION

### 1. Voluntary Attrition

This happens when workers voluntarily opt to leave the company.

Among the causes are:

- Offers of better jobs with more money or possibilities for advancement.
- Lack of advancement in career or satisfaction at work.
- Personal factors such as obligations related to family, health issues, or relocation.

It is frequently possible to stop voluntary attrition. High numbers typically point to issues with staff engagement, management, culture, or pay. To improve retention, HR must spot trends and deal with underlying issues.

### 2. Involuntary Attrition

This is what occurs when a company chooses to fire an employee.

Among the causes are:

- Termination for misbehaviour or poor performance.
- Layoffs brought on by cost-cutting, downsizing, or restructuring.

If it is not handled effectively, forced attrition may influence morale and is a reflection of the company's decisions. Effective processes and honest interactions are essential for preserving employee trust.

### 3. Healthy Attrition

This happens when workers who are disengaged, underperforming, or not aligned with the company leave on their own volition or are fired, this is known as healthy attrition.

Among the causes are:

- Despite coaching and support, employees routinely perform below expectations.
- People are hesitant to change, which hinders the company's ability to develop and expand.
- Some workers don't share the company's goals or are culturally incompatible.
- Teamwork can be impacted by instances of toxic behaviour or unfavourable attitudes.

Through the elimination of negative factors, the creation of space for dedicated workers, and the reinforcement of corporate principles and accountability, healthy attrition improves performance and morale.

#### 4. *Unhealthy Attrition*

This is the loss of key personnel, highly qualified experts, or outstanding individuals who are essential to the success of the company.

Among the causes are:

- Absence of prospects for advancement or recognition.
- Poisonous atmosphere at work or inadequate management.
- Goals of employees and organizational direction are not in line.

In particular, unhealthy attrition is harmful. High performers departure has an impact on morale and production and may even cause other workers to quit.

#### 5. *Retirement Attrition*

When workers leave their jobs upon retirement, usually after an age limit or a lengthy career, this is known as retirement attrition.

Among the causes are:

- Reaching the age of retirement, which is often between 58 and 65 based on legislation or corporate policy.
- Health concerns or a wish to lessen stress at work.
- Personal objectives like engaging in hobbies or spending more time with family.

## V. SYSTEM METHODOLOGY

### 1. *Data Collection*

This study's dataset was taken from the IBM HR Analytics Employee Attrition & Performance Dataset on Kaggle. It contains information about employees such as:

- Job Role, Department, Years at Company, Monthly Income, Job Satisfaction, etc.
- Demographic data like age, gender, marital status.
- Various performance metrics such as performance rating, attrition status.

Total there were 32 attributes (i.e., columns) and 1470 rows in this dataset. The target variable is "Attrition" with the classes values as: 'Yes', 'No'.

### 2. *Data Preprocessing*

Complete examination of the dataset was conducted by:

- Handling Missing Values: Missing value rows were eliminated and the median was used to fill in numerical missing values.
- Encoding Categorical values: Label encoding was used for the conversion of binary labels. To convert multiclass categorical variables, one-hot encoding was used.
- Feature Scaling: Label encoding was used for the conversion of binary labels. StandardScaler is normalization was implemented for scaling the features prior to supplying data to deep learning models.

### 3. *Exploratory Data Analysis*

- Disproportion between attrition classes was shown via count plots.
- Boxplots displayed the age distribution of departing and remaining employees.
- A feature correlation heatmap shed light on the connections between the variables.

### 4. *Models Implemented*

Machine Learning Models:

- Logistic Regression
- Random Forest Classifier
- Decision Tree Classifier
- K-Nearest Neighbours

Deep Learning Models:

- Artificial Neural Network
- Recurrent Neural Network

## VI. EXPERIMENTS & RESULTS

### 1. *Model Training and Evaluation*

Each model was trained using 80% of the data and 20% was reserved for testing.

Performance Metrics used were:

- Accuracy: Calculates the percentage of all accurate guesses among all predictions.

- Precision: Calculates how many of the positively predicted cases were actually positive.
- Recall: Calculates how many of the actual positive cases were correctly identified by the model.
- F1-score: The harmonic mean of precision and recall, balancing both false positives and false negatives.

## 2. Model Performance Comparison

The "Model Performance Comparison" bar chart compares several models, including Logistic Regression, Random Forest, Decision Tree, KNN, ANN and RNN, using four important metrics: Accuracy, Precision, Recall and F1-Score.

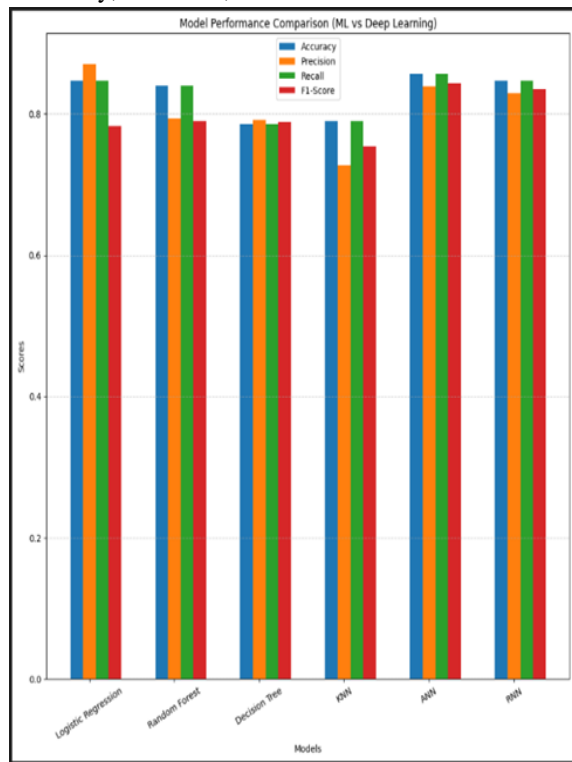


Fig: Model Performance Analysis

- Logistic Regression: In terms of accuracy and recall, logistic regression performed well overall, demonstrating its dependability in forecasting employee attrition.
- Random Forest: Random Forest's robustness and capacity to manage intricate patterns were demonstrated by the minor improvement in recall and F1-score compared to the majority of other models.

- Decision Tree and KNN: Despite their moderate performance, Decision Tree and KNN trailed somewhat in precision, indicating that they made a few more false positive predictions.
- Artificial Neural Network (ANN): The Artificial Neural Network's (ANN) ability to properly learn nonlinear relationships was demonstrated by its highly balanced performance across all criteria.
- Recurrent Neural Network (RNN): It also performed well, especially in recall, making it suitable for sequence-based patterns, even though the dataset wasn't strongly sequential.

## 3. Feature Importance

The top 10 most important features influencing employee attrition in this study found to be:

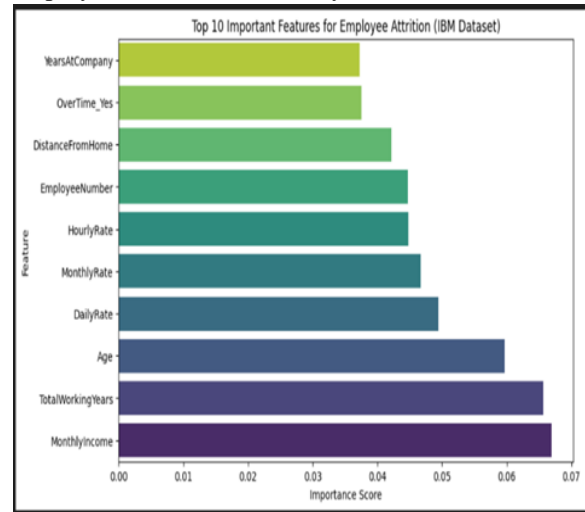


Fig: Top 10 Important Features for Employee Attrition

In the above graph x-axis represents the importance score which means how influencing this feature in the prediction model and y-axis represents the features that affects the employee attrition in any company. From this graph it is pretty much clear that the most 3 influential features are:

- Monthly Income*: It has the highest importance score which is above 0.066 which suggests that salary plays a very important role in most of the attrition decision. It is very clear from the study that people with lower income have the highest rate to leave the job.
- Total Working Years*: It has the second highest importance score which is nearly 0.065 which suggests that people with lesser experience are more likely to switch their jobs for better opportunities and career growth or they are at higher risk of leaving.

iii) *Age*: Career happiness and progress might be correlated with age. Younger workers could be more irrational or discontented.

Moderately Important Features are:

i) *Daily Rate, Monthly Rate, Hourly Rate*: A variety of pay-related factors are also rather significant. These could be a reflection of differences in pay plans.

Three Least influential features are:

i) *Distance From Home*: Employee turnover can increase due to stress and exhaustion caused by longer commutes. Additionally, it may lessen their bond with their job. It might be possible to lower attrition by offering remote or hybrid work options.

ii) *Overtime- Yes*: Regular overtime frequently results in poor work-life balance and burnout. If poorly managed, workers could feel overburdened and look for positions with more flexible schedules.

iii) *Years At Company*: An extended tenure may imply stagnation, whereas a brief tenure may show a lack of engagement. Attrition hazards can be present at either extreme if they are not well maintained.

#### 4. SHAP Analysis:

SHAP ideals provide more profound understanding:

Features on the Y-axis: Each row denotes a feature that was employed in the model, arranged in order of significance (from most to least effective).

A feature's X-axis (SHAP value) indicates how much it adds to the prediction's rise or fall.

The prediction is pushed toward attrition (or the positive class) when the SHAP value is positive.

A negative SHAP value pushes the prediction toward the negative class or away from attrition.

Variation in colour (feature value):

i) Red: High feature value

ii) Blue: Low feature value

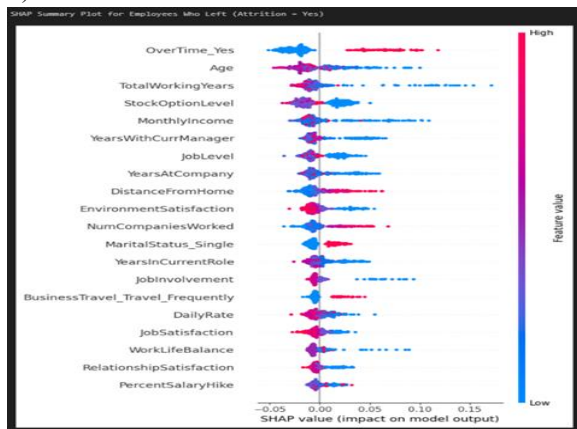


Fig: SHAP Summary for Attrition= YES

1. *Overtime-Yes*: Overtime workers (red) are more likely to experience attrition (positive SHAP values). Those who aren't working overtime (blue) usually stick around.

2. *Age*: Younger workers (blue) have a higher turnover rate. Negative SHAP values indicate a higher likelihood of retention for older workers (red).

3. *Total Years of Work*: Attrition is higher when total experience (blue) is lower. Attrition is decreased by more seasoned workers (red).

4. *Level of Stock Option*: Lower attrition is associated with higher stock options (red). The danger of quitting is increased by low or no stock options.

5. *Monthly Salary*: Attrition is higher among those with lower incomes (blue). Retention is encouraged by higher salary (red).

6. *Relationship satisfaction, work satisfaction and environment satisfaction*: Higher attrition is linked to lower satisfaction levels.

Summarizing the above SHAP Analysis:

Features	High Value Effect	Low Value Effect
Overtime-Yes	Increases Attrition	Reduces Attrition
Age	Reduces Attrition	Increases Attrition
Total Working Hours	Reduces Attrition	Increases Attrition
Monthly Income	Reduces Attrition	Increases Attrition
Environmental Satisfaction	Reduces Attrition	Increases Attrition
Job Satisfaction	Reduces Attrition	Increases Attrition
Work Life Balance	Reduces Attrition	Increases Attrition

Table: SHAP Analysis Summary

## VII. CONCLUSION

Using the IBM HR Analytics dataset, this study effectively forecasted employee attrition using both machine learning and deep learning models. Based on accuracy, precision, recall and F1-score, the Artificial Neural Network (ANN) performed the best overall out of all the models.

Age, Overtime, Total Working Years and Monthly Income were identified via feature importance analysis as the main determinants of attrition. These results imply that work-life balance and financial satisfaction are important factors in employee retention.

In conclusion, HR departments may identify at-risk workers and put plans in place to increase workforce stability and retention by using data-driven techniques.

## REFERENCES

- [1] Kishori Singh, Reetu Singh (August 2019)- “A Study on Employee Attrition: Effects and Causes”. International Journal of Research in Engineering, Science and Management. DOI: [https://www.ijresm.com/Vol.2\\_2019/Vol2\\_Iss8\\_August19/IJRESM\\_V2\\_I8\\_48.pdf](https://www.ijresm.com/Vol.2_2019/Vol2_Iss8_August19/IJRESM_V2_I8_48.pdf)
- [2] Mr. M. Lakshmanan, Mrs. R. Rogini kala (December 2024)- “Impact of Employee Attrition Rate in Industry”. International Journal of Progressive Research in Engineering Management and Science (ijprems).DOI:[https://www.ijprems.com/uploadedfiles/paper/issue\\_12\\_december\\_2024/37943/final/fin\\_ijprems1735668884.pdf](https://www.ijprems.com/uploadedfiles/paper/issue_12_december_2024/37943/final/fin_ijprems1735668884.pdf)
- [3] Shahin Manafi Varkiani, Francesco Pattarin, Tommaso Fabbri, Gualtiero Fantoni (13 January, 2025)- “Predicting Employee Attrition and Explaining its Determinants”. Expert Systems with Applications Journal (ELSEVIER). DOI:<https://www.sciencedirect.com/science/article/pii/S0957417425001976>
- [4] Dr. Sankararaman, Dr. Ragothaman, Balakrishnan, Divya Barathy, Dineshsurya (2023) – “A study on Understanding Employee Attrition- Causes and Consequences in the workplace”. European Economic Letters.
- [5] Rupa Chatterjee Das, D. A. (2020, October)- “Conceptualizing the Importance of HR Analytics in Attrition Reduction”. RPS SCIENCE HUB.
- [6] Annabelle Smyth -" Why Employee Retention Should Be a Top Priority". Social Hire.
- [7] Amit Gautam (2021). "Employee Engagement for the New Normal: The L&D Perspective". Invoice.