ISSN 2277-2685
IJESR/Apr-Jun. 2024/ Vol-14/Issue-2/204-209
Mr. A.Chandra Sekhar *et. al.,* / International Journal of Engineering & Science Research

PREDICTING EMPLOYEE ATTRITION RATE USING MACHINE LEARNING

1. Mr. A.Chandra Sekhar, M. Tech., (Ph.D) , Associate Professor, Department of Computer Science and Engineering (Artificial Intelligence), Gates Institute of technology, Gooty, Andhra Pradesh
chandra.ambarapu@gmail.com
2. Japthi Suraj,Student, Department of Computer Science and Engineering (Artificial Intelligence), Gates Institute of Technology-Gooty, Andhra Pradesh
japthisuraj6@gmail.com
3. Badvel Mohammad Avais, Student, Department of Computer Science and Engineering (Artificial Intelligence), Gates Institute of technology-Gooty, Andhra Pradesh
badvelavais18@gmail.com
4. B. Bharath Kumar Reddy, Student , Department of Computer Science and Engineering (Artificial Intelligence), Gates Institute of technology-Gooty, Andhra Pradesh
battalabharathkumarreddy143@gmail.com
5. U. Abinay , Student, Department of Computer Science and Engineering (Artificial Intelligence), Gates Institute of technology-Gooty, Andhra Pradesh
abhinani90142@gmail.com
6. N.Venu Gopal , Student, Department of Computer Science and Engineering (Artificial Intelligence), Gates Institute of technology-Gooty, Andhra Pradesh
venunesari@gmail.com

Abstract: Employees are an organization's most valuable resource. In any case, if they left their employment unexpectedly, it could result in a significant financial loss for any company. Because new hiring costs money and time, as well as requiring newly employed personnel to put in some effort to make the company productive. Following that, we aim to construct a model that would predict employee attrition rate based on HR analytics dataset in this research. To determine the reasons for employee attrition and reasons for leaving the organisation, a study called "Predicting Employee Attrition and Reasons for Leaving the Organization" was conducted. Why do the best and most experienced people leave the organisation, and how can you predict which valued personnel are likely to leave along these lines in order to figure out where the company is falling behind. This concept can be used by firms' Human Resource departments to develop effective strategies for retaining key executives before they start looking for new jobs, such as by increasing their remuneration.

I. INTRODUCTION

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy n company strength. Once these things r satisfied, ten next steps are to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration r taken into account for developing the proposed system.

Companies invest a lot of time and resources in employee recruiting and training, according to their strategic needs. Therefore, the employees (to a greater or lesser extent) represent a real investment for organisations. When an employee leaves the company, the organisation is not only losing a valuable employee, but also the resources, specifically money and HR staff effort, that were invested recruiting and selecting those employees and training them for their related tasks. Consequently, the organisation must continuously invest in recruiting, training and developing new staff to fill vacant job positions. Training a new employee is a long and costly process and it is of full interest of the company to control and decrease the employee attrition rate: attrition is defined as an employee resigning or retiring from a company. Moreover, satisfied, highly motivated and loyal employees form the core of a company and also have an impact on the productivity of an organisation. In the literature, some authors suggest retaining only happy and motivated employees as they tend to be more creative, productive and perform better, which in the end generates and sustains improved firm performance. As job dissatisfaction is shown in the economic literature as a good predictor of turnover intention job satisfaction data are powerful predictors of both separations and resignations, even controlling for wages, hours and standard demographic and job variables.

II. Literature Survey

we perform an analysis of the reasons or motivations that push an employee to leave the company and consequently allow the HR

department to take timely appropriate countermeasures such as improving the work environment or production incentives. Starting from the dataset, we identify the main factors related to the employee's attrition and we propose a real classification, based on the statistical evaluation of the data. The application of classification algorithms can support the HR management by allowing the adoption of staff management support tools in the company. The obtained model for the prediction of employees' attrition is tested on a real dataset provided by IBM analytics, which includes 35 features and about 1500 samples. By analysing the correlations in the heatmap of 35 features, we derive the characteristics that have high correlations related to the reasons that an employee leaves the company.

a. ***Logistic Regression***: Logistic regression is one of the most widely used machine learning techniques for Employee attrition rate prediction. This method has been applied to various datasets and has been found to be effective in detecting the presence of Employee attrition rate.
b. ***Decision Trees*****:** Decision trees are another popular machine learning technique that has been used for Employee attrition rate prediction. The method works by creating a tree-like model that predicts the outcome based on certain attributes. This technique has been found to be effective in predicting Employee attrition rate, especially when combined with other machine learning techniques.
c. **Gaussian Naive Bayes*:*** Naïve Bayes is part of a family of generative learning algorithms, meaning that it seeks to model the distribution of inputs of a given class or category. Unlike discriminative classifiers, like logistic regression, it does not learn which features are most important to differentiate between classes.
d. ***Support Vector Machines (SVMs):*** SVMs are another machine learning technique that has been applied to Employee attrition rate prediction. The method works by finding the best boundary between the different classes of data and has been found to be effective in predicting Employee attrition rate.

e. ***Random Forests:*** Random forests are an ensemble learning technique that combines multiple decision trees to make a prediction. This technique has been found to be effective in Employee attrition rate prediction, especially when dealing with high-dimensional data.

III. Existing System

Attrition is the term used to describe the process of reducing the number of employees. These wordings can be used to analyze manpower and other estimates that are important for manpower planning. Turnover occurs as a result of a variety of work actions, such as release, termination, abandonment, or, on the other side, occupation surrender. Attrition occurs when a worker leaves or when the organization eliminates his position. The difference between the two is that when an employee leaves, the company looks for someone to take their place. Attrition occurs when a business leaves an opportunity unfulfilled or clears out business tasks.

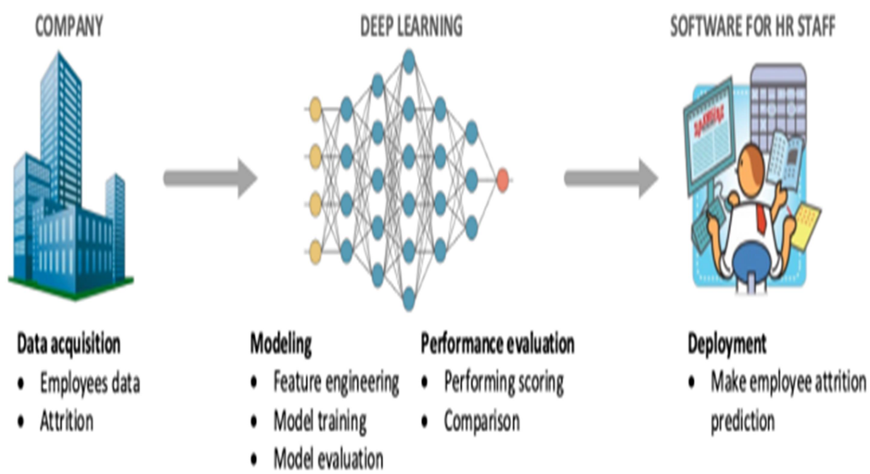

FIG1: Existing System

IV. Proposed system

Predicting staff turnover in a company allows management to act more quickly by updating internal policies and initiatives. Where talented employees who are on the verge of leaving might be offered a few suggestions, such as a wage raise or suitable training, to lessen their likelihood of leaving. Machine learning algorithms can assist firms in predicting staff attrition. Analysts can construct and prepare a machine learning model that can anticipate the personnel who are leaving the firm using previous data kept in human resources (HR) departments.

A proposed system for predicting the onset of Employee attrition rate using machine learning could include the following components:

a) **Data collection:** The first step would be to collect relevant data on patients, such as demographic information, medical history, lifestyle habits, and physiological measurements. This data would be used as input for the machine learning algorithm.
b) **Feature engineering:** In this step, the collected data would be pre-processed and relevant features would be selected and transformed to improve the performance of the machine learning algorithm.
c) **Model selection and training:** The next step would be to select a suitable machine learning algorithm, such as decision trees, random forests, support vector machines, or deep learning methods, and train the model on the collected data. This step would involve splitting the data into training and testing sets, and tuning the parameters of the model to optimize its performance.
d) **Model evaluation:** Once the model has been trained, it would be evaluated on the testing set to determine its accuracy in predicting the onset of Employee attrition rate .
e) **Deployment:** The final step would be to deploy the model in a real-world setting, such as a hospital or healthcare provider's office. The model would be used to analyze patient data and make predictions about their risk for Employee attrition rate .

V. CONCLUSION

we applied some machine learning techniques in order to identify the factors that may contribute to an employee leaving the company and, above all, to predict the likelihood of individual employees leaving the company. First, we assess statistically the data and then we classified them. The dataset was processed, dividing it into the training phase and the test phase, guaranteeing the same distribution of the target variable (through the holdout technique). We selected various classification algorithms and, for each of them, we carried out the training and validation phases. To evaluate the algorithm's performance, the predicted results were collected and fed into the respective confusion matrices. From these it was possible to calculate the basic metrics necessary for an overall evaluation (precision, recall, accuracy, f1 score, ROC curve, AUC, etc.) and to identify the most suitable classifier to predict whether an employee was likely to leave the company. The algorithm that produced the best results for the available dataset was the Gaussian Naïve Bayes classifier: it revealed the best recall rate (0.54), a metric that measures the ability of a classifier to find all the positive instances, and achieved an overall false negative rate equal to 4.5% of the total observations. Results obtained by the proposed automatic predictor demonstrate that the main attrition variables are monthly income, age, overtime, distance from home. The results obtained from the data analysis represent a starting point in the development of increasingly efficient employee attrition classifiers. The use of more numerous datasets or simply to update it periodically, the application of feature engineering to identify new significant characteristics from the dataset and the availability of additional information on employees would improve the overall knowledge of the reasons why employees leave their companies and, consequently, increase the time available to personnel departments to assess and plan the tasks required to mitigate this ris

VI. References

[1. Cockburn, I.; Henderson, R.; Stern, S. The Impact of Artificial Intelligence on Innovation. In The Economics of Artificial Intelligence: An Agenda; University of Chicago Press: Chicago, IL, USA, 2019; pp. 115–146.

2. Jarrahi, M. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Bus. Horiz. 2018, 61, 577–586. [CrossRef]

3. Yanqing, D.; Edwards, J.; Dwivedi, Y. Artificial intelligence for decision making in the era of Big Data. Int. J. Inf. Manag. 2019, 48, 63–71.

4. Paschek, D.; Luminosu, C.; Dra, A. Automated business process management-in times of digital transformation using machine learning or artificial intelligence. In MATEC Web of Conferences; EDP Sciences: Les Ulis, France, 2017; Volume 121.

5. Varian, H. Artificial Intelligence, Economics, and Industrial Organization; National Bureau of Economic Research: Cambridge, MA, USA, 2018.

6. Vardarlier, P.; Zafer, C. Use of Artificial Intelligence as Business Strategy in Recruitment Process and Social Perspective. In Digital Business Strategies in Blockchain Ecosystems; Springer: Berlin/Heidelberg, Germany, 2019; pp. 355–373.

7. Gupta, P.; Fernandes, S.; Manish, J. Automation in Recruitment: A New Frontier. J. Inf. Technol. Teach. Cases 2018, 8, 118–125. [CrossRef]

8. Geetha, R.; Bhanu Sree Reddy, D. Recruitment through artificial intelligence: A conceptual study. Int. J. Mech. Eng. Technol. 2018, 9, 63–70.

9. Syam, N.; Sharma, A. Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. Ind. Mark. Manag. 2018, 69, 135–146. [CrossRef]

10. Mishra, S.; Lama, D.; Pal, Y. Human Resource Predictive Analytics (HRPA) For HR Management in Organizations. Int. J. Sci. Technol. Res. 2016, 5, 33–35.

11. Jain, N.; Maitri. Big Data and Predictive Analytics: A Facilitator for Talent Management. In Data Science Landscape; Springer: Singapore, 2018; pp. 199–204.

12. Boushey, H.; Glynn, S.J. There Are Significant Business Costs to Replacing Employees. Cent. Am. Prog. 2012, 16, 1–9.

13. Martin, L. How to retain motivated employees in their jobs? Econ. Ind. Democr. 2018, 34, 25–41. [CrossRef]

14. Wood, S.; Van Veldhoven, M.; Croon, M.; de Menezes, L.M. Enriched job design, high involvement management and organizational performance: The mediating roles of job satisfaction and wellbeing. Hum. Relat. 2012, 65, 419–446. [CrossRef]

15. Zelenski, J.M.; Murphy, S.A.; Jenkins, D.A. The happy-productive worker thesis revisited. J. Happiness Stud. 2008, 9, 521–537. [CrossRef]

16. Clark, A.E. What really matters in a job? Hedonic measurement using quit data. Labour Econ. 2001, 8, 223–242. [CrossRef]

17. Clark, A.E.; Georgellis, Y.; Sanfey, P. Job satisfaction, wage changes, and quits: Evidence from Germany. Res. Labor Econ. 1998, 17, 95–12118. Delfgaauw, J. The effect of job satisfaction on job search: Not just whether, but also where. Labour Econ. 2007, 14, 299–317. [CrossRef]

19. Green, F. Well-being, job satisfaction and labour mobility. Labour Econ. 2010, 17, 897–903. [CrossRef]

20. Kristensen, N.; Westergaard-Nielsen, N. Job satisfaction and quits – which job characteristics matters most? Dan. Econ. J. 2006, 144, 230–249.

21. Marchington, M.; Wilkinson, A.; Donnelly, R.; Kynighou, A. Human Resource Management at Work; Kogan Page Publishers: London, UK, 2016.

22. Van Reenen, J. Human resource management and productivity. In Handbook of Labor Economics; Elsevier: Amsterdam, The Netherlands, 2011.

23. Deepak, K.D.; Guthrie, J.; Wright, P. Human Resource Management and Labor Productivity: Does Industry Matter? Acad. Manag. J. 2005, 48, 135–145.

24. Gordini, N.; Veglio, V. Customers churn prediction and marketing retention strategies. An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry. Ind. Mark. Manag. 2016, 62, 100–107. [CrossRef]

25. Keramati, A.; Jafari-Marandi, R.; Aliannejadi, M.; Ahmadian, I.; Mozaffari, M.; Abbasi, U. Improved churn prediction in telecommunication industry using data mining techniques. Appl. Soft Comput. 2014, 24, 994–1012. [CrossRef]