**EMPLOYEE CHURN**

**A PROJECT REPORT**

*Submitted by*

Ankit Kumar Pradhan

*in partial fulfilment for the award of the*

*degree of*

BACHELOR OF TECHNOLOGY

*In*

COMPUTER SCIENCE ENGINEERING



**CENTURION UNIVERSITY OF TECHNOLOGY & MANAGEMENT** BHUBANESWAR– 752050

**9 Ocober, 2021**

**1 |** P a g e

SPECIMEN CERTIFICATE

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING CENTURION UNIVERSITY OF TECHNOLOGY AND MANAGEMENT, BHUBANESWAR-752050

BONAFIDE CERTIFICATE

Certified that this project report **“EMPLOYEE CHURN”** is the bonafide work of **“Ankit Kumar Pradhan”** who carried out the project work under my supervision. This is to further certify to the best of my knowledge, that this project has not been carried out earlier in this institute and the university.

**SIGNATURE**

**(Signature of Supervisor)**

*Certified that the above-mentioned project has been duly carried out as per the norm of the college and statutes of the university*

**SIGNATURE**

**PROF. (DR.)** **SUJATA CHAKRAVARTY**

**HEAD OF THE DEPARTMENT**

DEPARMENTAL SEAL

**2 |** P a g e

**ACKNOWLEDGEMENTS**

It is our pleasure to be indebted to various people, who directly or indirectly contributed in the development of this work and who influenced my/our thinking, behavior and acts during the course of study.

I am thankful to Prof. Sujata Chakravarty for her support, cooperation, and motivation provided to me during for constant inspiration, presence and blessings.

I also extend our sincere appreciation to Prof. Sujata Chakravarty, Head of Department who provided his valuable suggestions and precious time in accomplishing our academic project.

**Name: Ankit Kumar Pradhan**

**Reg-No: 190301120005**

**3 |** P a g e

TABLE OF CONTENTS

1. CHAPTER-1 INTRODUCTION

2. CHAPTER-2 ABOUT THE PROJECT

3. CHAPTER -3 DATASET

4. CHAPTER-4 CONCLUSION

5. CHAPTER-5 REFERENCES

**ABSTRACT**

Employee churn is a major issue in companies as it affects the company cost and to replace the left employee with a new employee. So, predictive models are useful in this type of scenarios. Employee churn and employee attrition are closely related but they are not the same. Employee churn affects its overall revenue and brand image of the organization. Employee churn leads to customer dissatisfaction. New hiring will consume money as well as time and freshly hired employees take time to make the respective organization profitable. Correlation matrix plot is also done on our dataset to see which features are more related to our target variable. Some more visualization works have been done like comparing how many employees left the company, left vs. satisfaction level, left vs. promotion since last 5 years. Well- known classification algorithms including, Decision Tree, SVM, KNN have been applied. Different performance measures like precision, recall, F-measure and overall accuracy are taken into consideration. From the simulation results it can easily be concluded that decision tree performs better as compared to other models with the accuracy of 98.0677%.

1. **INTRODUCTION**

An employee would decide to join or leave an organization based on several reasons, for instance, work environment, work place, gender equity, pay equity etc. Others may consider personal reasons such as relocation due to family, maternity, health, conflict with the managers or colleagues in a team. Employee churn is a big issue for the organizations specially when trained, technical and key employees leave for a better opportunity in a competitor organization. It requires time, effort and results in financial loss to replace a trained employee. Therefore, we use the current and past employee data to analyze the common grounds for employee attrition. The employee churn prediction helps in identifying and solving the issues that results in attrition. We can use this information for possible retention of the current employees. In this study, we implement some of the well-known techniques of data classification namely, Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Artificial Neural Network (ANN) on the data set. The dataset includes 1499 records with 9 features including categorical and numeric features. Before implementing method, we calculated the correlation of the features in order to avoid features with high correlation. The results of this methods have been analyzed then by their accuracy, precision, recall, and F-measure values. Then, the method with best performance has been conducted. Finally, we implement a feature selection method to select the most important features of the dataset.

**2 About the Project**

Churn analysis and prediction is widely studied for customer churn, since it is a notorious issue and results in revenue loss. Employee churn is a similar problem for organization, however to predict employee churn is rather more complex than customer churn prediction. Employee churn leads to issues such as efforts and time to get the replacement and retraining, financial loss, customers dissatisfaction and many more. Therefore, for smooth running of an organization, the key is to retain its trained workforce. Employee churn can be categorized in two types; (1) voluntary, those who leave for their own reasons, and (2) involuntary, those who are released from their services by the organization. Usually companies focus on voluntary churn, where an employee would either leave for a better opportunity in terms of pay, benefits, work environment etc., or due to negative reasons at the present organization such as conflict with the supervisors, lack of opportunities for promotion, lack of interesting work and many more. In this study, we also focus on predict voluntary churn employees.

**3 DATASETS**

There are two datasets used in the project on is hr\_data in which we have 9 features, which are employee\_id, number\_project, average\_monthly\_hours, time\_spend\_company, Work\_accident, Left, promotion\_last\_5\_years, department, salary. Department and salary are categorical values. The other dataset is employee\_satisfaction\_evaluation, in that dataset we have 3 features which are EMPLOYEE#, satisfaction\_level, last\_evaluation. The main data frame has been created using both the datasets where employee\_id and EMPLOYEE# were primary keys.

**4. CONCLUSION**

Employee churn leads to customer dissatisfaction. In this project, hr\_data and employee\_satisfaction\_evaluation datasets have been considered. Different machine learning algorithms, Decision tree, KNN, SVM, and ANN have been used to find out which employee is going to quit the respective organization based on their working details and environments. Different performance measures like confusion matrix and overall accuracy have been used. From the experimental results, Decision tree is found to be better as compared to other models considered in this study. Thus, this model can be used for the real-life applications.

**5. REFERENCE**

[1] Yiğit, İbrahim Onuralp, and Hamed Shourabizadeh. "An approach for predicting employee churn by using data mining." *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*. IEEE, 2017.

[2] Saradhi, V. Vijaya, and Girish Keshav Palshikar. "Employee churn prediction." *Expert Systems with Applications* 38.3 (2011): 1999-2006.

[3] Sisodia, Dilip Singh, Somdutta Vishwakarma, and Abinash Pujahari. "Evaluation of machine learning models for employee churn prediction." *2017 International Conference on Inventive Computing and Informatics (ICICI)*. IEEE, 2017

[4] Dolatabadi, Sepideh Hassankhani, and Farshid Keynia. "Designing of customer and employee churn prediction model based on data mining method and neural predictor." *2017 2nd International Conference on Computer and Communication Systems (ICCCS)*. IEEE, 2017.

[5] Alamsyah, Andry, and Nisrina Salma. "A comparative study of employee churn prediction model." *2018 4th International Conference on Science and Technology (ICST)*. IEEE, 2018.

[6] Jain, Nishant, Abhinav Tomar, and Prasanta K. Jana. "A novel scheme for employee churn problem using multi-attribute decision making approach and machine learning." *Journal of Intelligent Information Systems* 56.2 (2021): 279-302.

[7] Srivastava, Praveen Ranjan, and Prajwal Eachempati. "Intelligent Employee Retention System for Attrition Rate Analysis and Churn Prediction: An Ensemble Machine Learning and Multi-Criteria Decision-Making Approach." *Journal of Global Information Management (JGIM)* 29.6 (2021): 1-29.

[8] Ajit, Pankaj. "Prediction of employee turnover in organizations using machine learning algorithms." *algorithms* 4.5 (2016): C5.

[9] Jain, Himani, Garima Yadav, and R. Manoov. "Churn Prediction and Retention in Banking, Telecom and IT Sectors Using Machine Learning Techniques." *Advances in Machine Learning and Computational Intelligence*. Springer, Singapore, 2021. 137-156.

[10] Khera, Shikha N., and Divya. "Predictive modelling of employee turnover in Indian IT industry using machine learning techniques." *Vision* 23.1 (2018): 12-21.