

Table of Contents

INTRODUCTION	6
OVERVIEW.....	6
PURPOSE.....	6
LITERATURE ANALYSIS.....	6
EXISTINGPROBLEM.....	6
PROPOSEDSOLUTION.....	8
THEORITICAL ANALYSIS.....	9
BLOCK DIAGRAM	9
 HARDWARE/SOFTWAREDESIGNING	10
 EXPERIMENTAL INVESTIGATION.....	10
FLOWCHART	15
RESULT	15
 ADVANTAGES & DISADVANTAGES	17
 APPLICATIONS	18
CONCLUSION	18
FUTURESCOPE.....	19
BIBIOGRAPHY	19

1.INTRODUCTION

OVERVIEW

The problem of employee turnover has shot to prominence in organizations because of its negative impacts on issues ranging from work place morale and productivity, to disruptions in project continuity and to long term growth strategies.

One way organizations deal with this problem is by predicting the risk of attrition of employees using machine learning techniques thus giving organizations leaders and Human Resources (HR) the foresight to take pro-active action for retention or plan for succession.

However, the machine learning techniques historically used to solve this problem fail to account for the noise in the data in most HR Information Systems (HRIS). Most organizations have not prioritized investments in efficient HRIS solutions that would capture an employee's data during his/her tenure.

One of the major factors is the limited understanding of benefits and cost. It is still difficult to measure the return of investment in HRIS . This leads to noise in the data, which in turn attenuates the generalization capability of these algorithms

PURPOSE:

Employee Promotion plays a big role in Employee Satisfaction. It aids in employee engagement, boosts morale, reduces absenteeism, and ultimately in productivity. Employee Promotion is also a helpful tool for reducing attrition retention

2. LITERATURE SURVEY

EXISTING PROBLEM:

To overcome this problem, organizations are now taking support via machine learning techniques to predict the employee turnover.

With high precision in prediction, organizations can take necessary actions at due course of time for retention or succession of employees.

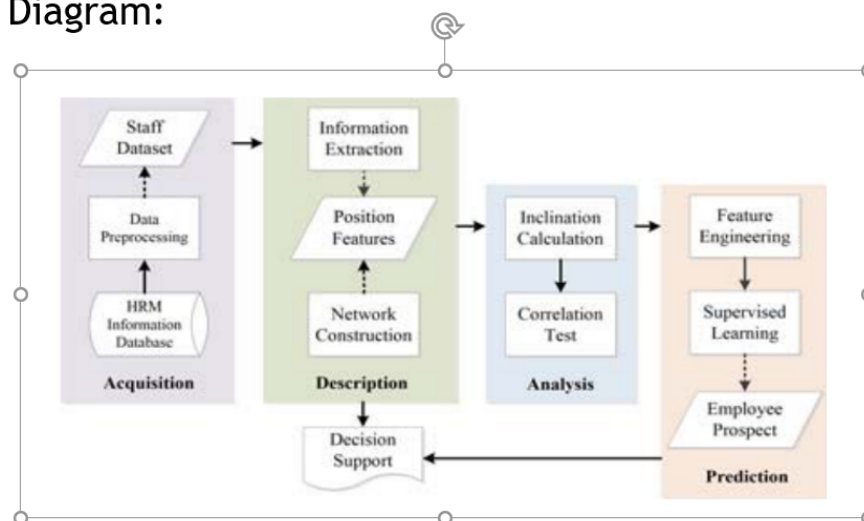
Most of the data comes from basic HR based database systems, which are not highly efficient in prediction and modeling and these models are not very accurate in data models and cannot assist the organizations to take successful decisions. The primary objective of this research paper is to predict employee attrition i.e. whether the employee is planning to leave or continue to work within the organization

Data collection

Collecting data allows you to capture a record of past events so that we can use data analysis to find recurring patterns. From those patterns, you build predictive models using machine learning algorithms that look for trends and predict future changes.

Block Diagram:

Block Diagram:



Hardware Requirement:

System : Pentium i3 Processor.

Hard Disk : 500 GB.

Monitor : 15'' LED

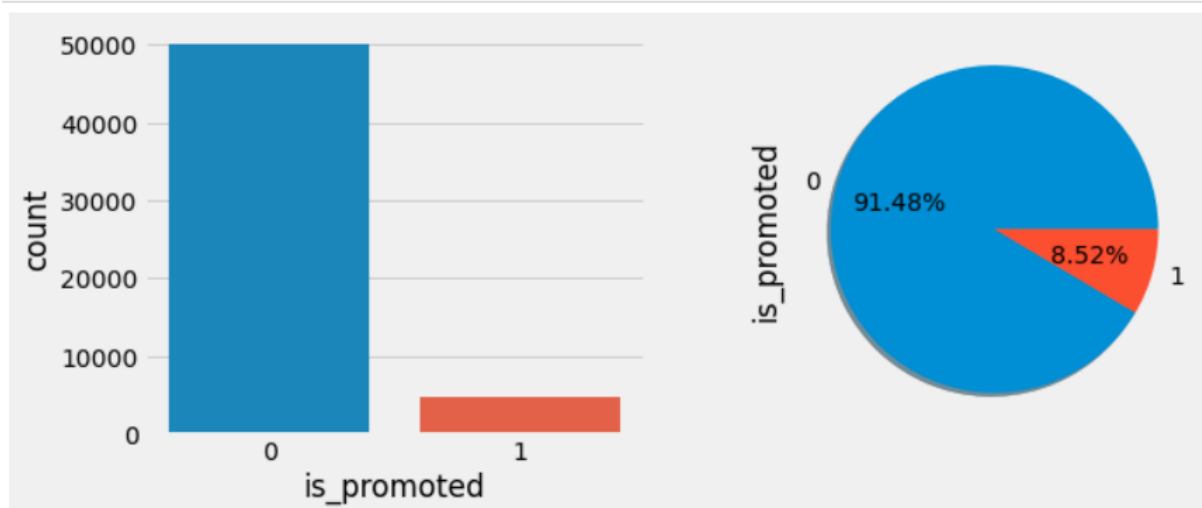
Input Devices : Keyboard, Mouse

Ram : 2 GB

Software Requirement:

Operating system : Windows 10.

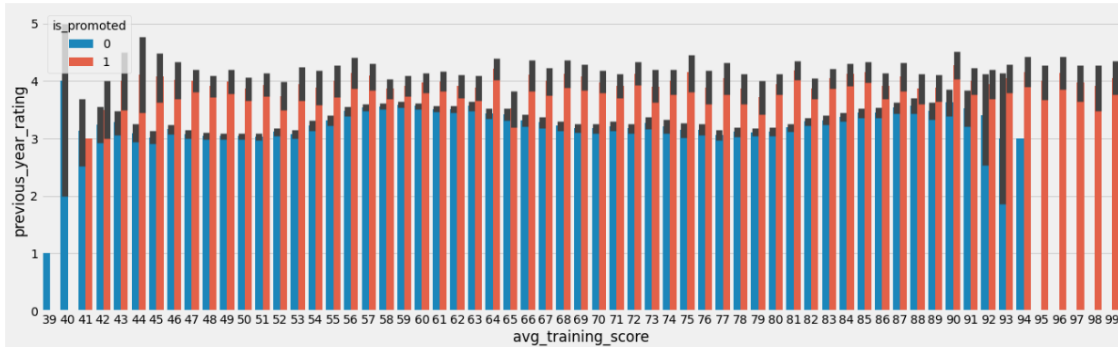
Coding Language : Python



MULTIVARIATE

Multivariate analysis

```
In [9]: """ From the below bar plot, we came to know that employee with training score > 95 & previous year rating > 3 got promoted. """  
plt.figure(figsize=(20,6))  
sns.barplot(df['avg_training_score'],df['previous_year_rating'],df['is_promoted'])  
  
Out[9]: <AxesSubplot:xlabel='avg_training_score', ylabel='previous_year_rating'>
```



Advantage

increasing customer traffic.

The more you promote your brand, the more will the customers know about you and your company and the more will they be interested in your products.

Promotion can be done even by giving out free samples which work wonders for customers

DISADVANTAGE:

No new or fresh ideas are brought into the organisation.

The job advertised may require skills not currently available within the organisation.

Promotion of an internal employee could cause resentment

amongst other employees, who may feel they deserve the post more than the promoted employee.

TESTING MODEL

In this module we test the trained machine learning model using the test dataset

Quality assurance is required to make sure that the software system works according to the requirements. Were all the features implemented as agreed? Does the program behave as expected? All the parameters that you test the program against should be stated in the technical specification document.

Moreover, software testing has the power to point out all the defects and flaws during development. You don't want your clients to encounter bugs after the software is released and come to you waving their fists. Different kinds of testing allow us to catch bugs that are visible only during runtime.

increasing customer traffic.

The more you promote your brand, the more will the customers know about you and your company and the more will they be interested in your products.

Promotion can be done even by giving out free samples which work wonders for customers

DATA COLLECTION

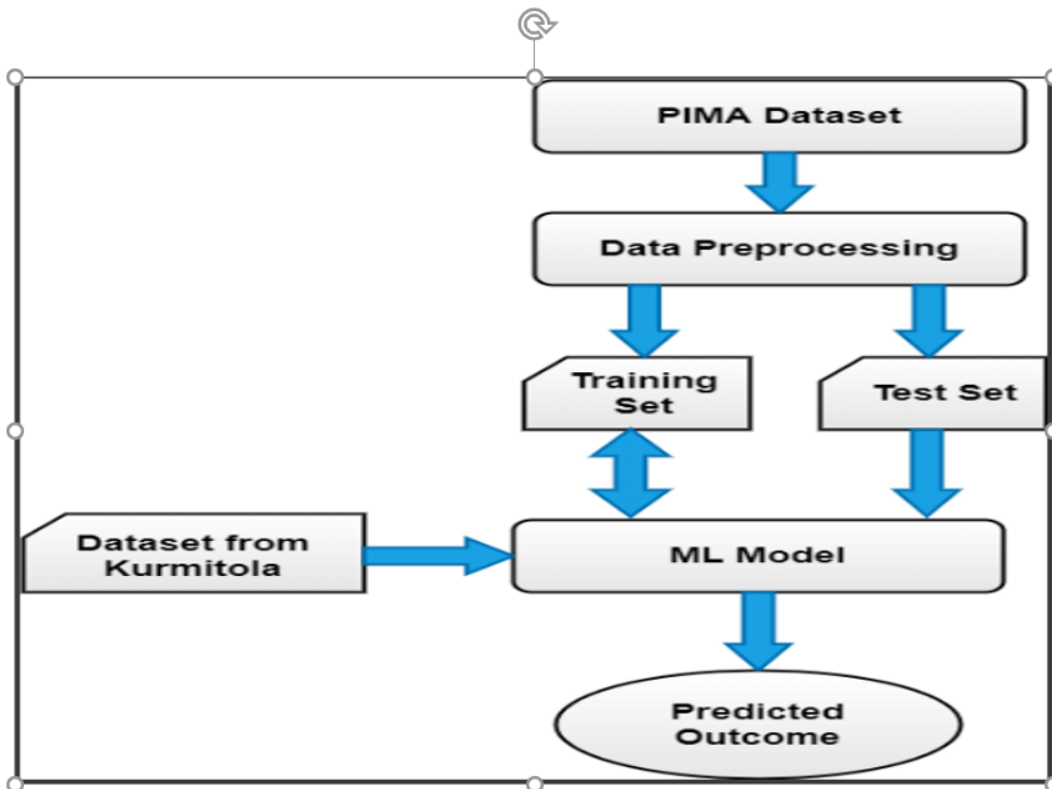
Collecting data allows you to capture a record of past events so that we can use data analysis to find recurring patterns. From those patterns, you build predictive models using machine learning algorithms that look for trends and predict future changes.

Predictive models are only as good as the data from which they are built, so good data collection practices are crucial to developing high-performing models.

The data need to be error-free (garbage in, garbage out) and contain relevant information for the task at hand. For example, a loan default model would not benefit from tiger population sizes but could benefit from gas prices over time.

In this module, we collect the employee promotion prediction data from kaggle dataset archives. This dataset contains the information of divorce in previous years.

5.FLOW CHART



6. RESULT AND CONCLUSION

The importance of predicting employee turnover in organizations and the application of machine learning in building turnover models was presented in this paper.

The key challenge of noise in the data from HRIS that compromises the accuracy of these predictive models was also highlighted. Data from the HRIS of a global retailer was used to compare the XGBoost classifier against six other supervised class that had been historically used to build turnover models. The results of this research demonstrate that the XGBoost classifier is a superior algorithm in terms of significantly higher accuracy, relatively low runtimes and efficient memory utilization for predicting turnover.

its regularization makes it a robust technique capable of handling the noise in the data from HRIS, as compared to the other classifiers, thus overcoming the key challenge in this domain.

Because of these reasons it is recommended to use XGBoost for accurately predicting employee turnover, thus enabling to take actions for retention or succession of employees.

REFERENCES:

1. [1] S. Jahan, "Human Resources Information System (HRIS): A Theoretical Perspective", Journal of Human Resource and Sustainability Studies, Vol.2 No.2, Article ID:46129, 2014.
2. [2] M. Stoval and N. Bontis, "Voluntary turnover: Knowledge management – Friend or foe?", Journal of Intellectual Capital, 3(3), 303-322, 2002.
3. [3] J. L. Cotton and J. M. Tuttle, "Employee turnover: A meta-analysis and review with implications for research", Academy of management Review, 11(1), 55-70, 1986.
4. [4] L. M. Finkelstein, K. M. Ryan and E.B. King, "What do the young (old) people think of me? Content and accuracy of age-based metastereotypes", European

Journal of Work and Organizational Psychology,
22(6), 633-657, 2013.

5. [5] B. Holtom, T. Mitchell, T. Lee, and M. Eberly,
“Turnover and retention research: A glance at the
past, a closer review of the present, and a venture
into the future”, Academy of Management Annals, 2:
231-274, 2008