EMPLOYEE ATTRITION PREDICTION

PRESENTED BY :

NIVEEN PRABHU

PARTHASARATHI S

NIRUBAMA P G

PREETHIPRIYADARSHNI P

PROJECT REPORT

1.INTRODUCTION :

1.1 OVERVIEW :

The project aims to provide the organization the overview of its employee and their attrition. Employee attrition is a challenging issue in the business world. It has a significant impact on the competitive strength of a company. As per the Report on Absenteeism, Labor turnover, Employment and labor cost, prepared by the Ministry of Labor and Employment; Labor Bureau, Employee attrition in the automobile sector in India is 40 percent. Employee attrition creates chronic problems in an organization. The impact of employee attrition has been realized by industrialists during the past couple of years.

1.2 PURPOSE :

Human resource is an important and a valuable asset of an organization. However, most of the organizations consider it as a cost center. Employees play a pivotal role in an organization because all the resources are wasted if they are not properly and efficiently utilized by the employees. It is the employee who can create a history by making an organization or by breaking an organization. Today in a highly competitive and dynamic business world the success of an

organization depends on the ability and stability of competitive human resource. Maintaining and retaining existing employees is the most challenging issue for any organization. The project serves to overcome the above addressed issue.

2.LITERATURE SURVEY :

2.1 EXISTING PROBLEM:

Attrition basically rises due to many factors like dissatisfaction with the company, insufficient salary, poor working environment, lack of motivation, poor relations with colleagues and boss, unhappiness of an employee and many other factors. The attrition level of employees indicates that something is wrong with the health and climate of an organization in terms of wages, working conditions, industrial relation, welfare facilities provided by the employer to the employees etc. Higher rate of attrition indicates employee unrest and lack of stability in the labour force which is not good for competitiveness, growth and development of an organization. The organization faces uncertain cost, disturbance in production and work atmosphere, cost of recruitment, selection, training and development and so on. These uncertainties raise a question regarding the reasons for attrition in automobile industries. The organisation views the employee attrition prediction process as time-consuming and expensive.

2.2 PROPOSED SOLUTION :

The proposed solution aims to predict the employee attrition inclusive of all the categories related to the organisation and the

employee. It helps the organisation an opportunity to discover its own employee, which is a key point in the development and growth of the organisation. It also offers a personalised service to overcome the issues faced between the management and the employee community.

3. EXPERIMENTAL INVESTIGATIONS :

Planning and project setup

- *Define the task and scope out requirements*
- *Determine project feasibility*
- *Discuss general model tradeoffs (accuracy vs speed)*
- *Set up project codebase*

Data collection and labeling

- *Define ground truth (create labeling documentation)*
- *Build data ingestion pipeline*
- *Validate quality of data*
- *Revisit Step 1 and ensure data is sufficient for the task*

Model exploration

- *Establish baselines for model performance*
- *Start with a simple model using initial data pipeline*
- *Overfit simple model to training data*
- *Stay nimble and try many parallel (isolated) ideas during early stages*

↳ *Find SoTA model for your problem domain (if available) and reproduce results, then apply to your dataset as a second baseline*

↳ *Revisit Step 1 and ensure feasibility*

↳ *Revisit Step 2 and ensure data quality is sufficient*

Model refinement

↳ *Perform model-specific optimizations (ie. hyperparameter tuning)*

↳ *Iteratively debug model as complexity is added*

↳ *Perform error analysis to uncover common failure modes*

↳ *Revisit Step 2 for targeted data collection of observed failures*

Testing and evaluating

↳ *Evaluate model on test distribution; understand differences between train and test set distributions.*

↳ *Revisit model evaluation metric; ensure that this metric drives desirable downstream user behavior.*

↳ *Write tests for:*

- *Input data pipeline*
- *Model inference functionality*
- *Model inference performance on validation data*
- *Explicit scenarios expected in production (model is evaluated on a curated set of observations)*

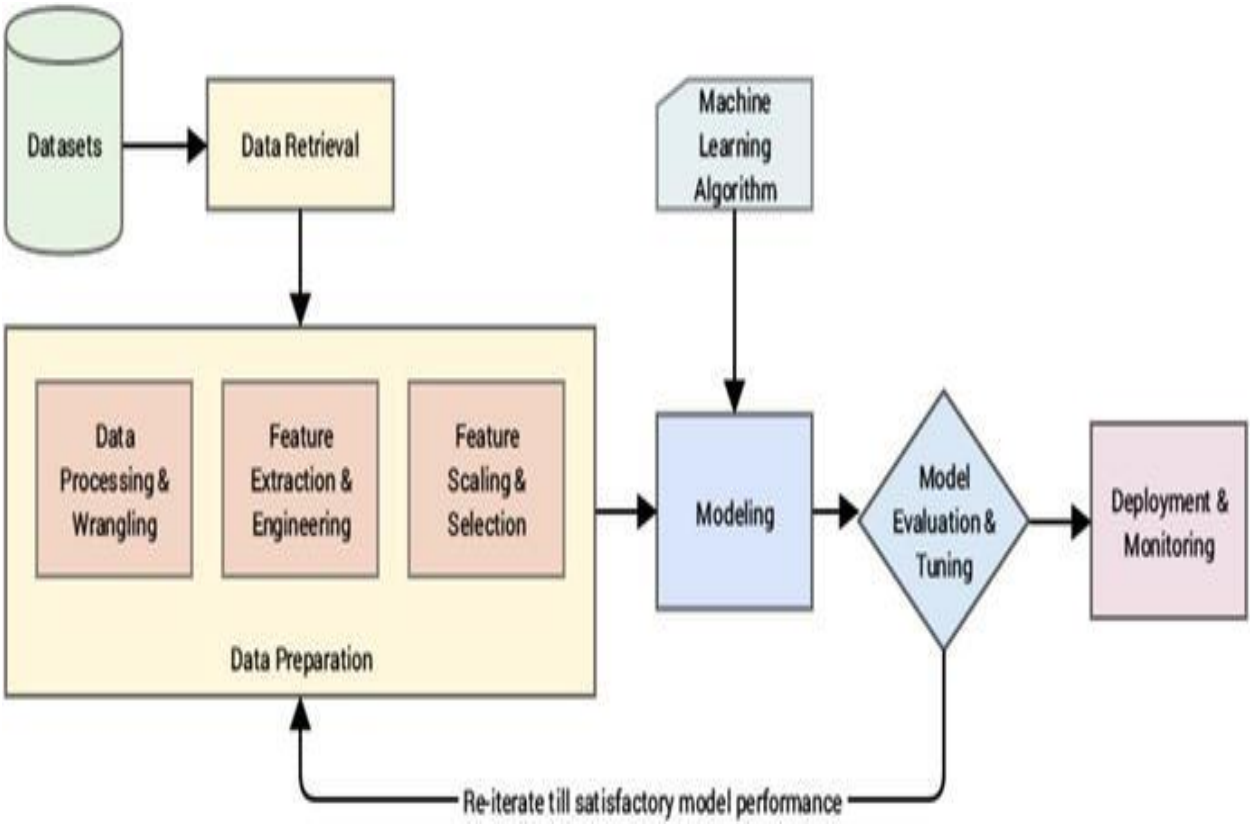
Model deployment

- *Expose model via a REST API*
- *Deploy new model to small subset of users to ensure everything goes smoothly, then roll out to all users*
- *Maintain the ability to roll back model to previous versions*
- *Monitor live data and model prediction distributions*

Application development

- *Includes web development or app development based on the user demands.*
- *Deployment in cloud is the best practice for machine learning related projects.*

4. FLOW CHART :



The above flow chart describes the general overview of machine learning projects. This model is followed universally for better performance . Initially the datasets are collected and obtained and data wrangling techniques are performed. The values to be binary encoded are done and feature scaling takes place here. After feature scaling, extraction and selection is carried out. Now efficient model for the given dataset is chosen and and model is created. The next step is to evaluate the model. After the evaluation process factors related to accuracy are tuned to obtain greater accuracy. Next step is to deploy the completed module as a web serivce or an application based on the user demand. Re-iteration is also done if the performance of the module is not satisfactory.

5.RESULT :

By using Machine learning model we are try to predict which valuable employees are probable to leave the organization subsequently so as to find the areas where the organization is lagging behind. This model can be used by the Human Resource departments of the organizations to form efficient strategies to retain the valuable employees before they start looking for new jobs like by providing a hike in their salary, offering promotions if necessary travel and stay abroad or start the hiring process.

EMPLOYEE ATTRITION PREDICTION

EDUCATION :

college

JOB INVOLVEMENT:

Very high

1

5644

1233

2

1

6

0

0

4

2

Outstanding

Predict

EMPLOYEE ATTRITION PREDICTION

EDUCATION :

JOB INVOLVEMENT:

JobLevel

DailyRate(USD)

MonthlyIncome(USD)

NoofCompaniesWorked

TotalWorkingYears

YearsAtCompany

YearsInCurrentRole

YearsSinceLastPromotion

YearsWithCurrentManager

TrainingTimesLastYear

PerformanceRating

Predict

EMPLOYEE Attrition YES

6.ADVANTAGES & DISADVANTAGES :

ADVANTAGES:

- **Higher manpower costs:**

There are times when employees stay with the organization for long, which might mean that they are getting top of their pay scale.This means that these employees are being paid a lot more than others who are doing similar job but are comparatively new in the process.This excessive manpower cost leads to financial burden which is generally not identified on a regular run .

- **New idea:**

Many a time when some people leave an organization they open gates for new talent and new ideas.Mostly employees who are in the organization get used to the working atmosphere and get complacent. This means lack of risk and definitely no new ideas.When they leave there is space for people who have high risk taking caliber and stop the firm from becoming stagnant.

- **Higher performance:**

There are employees who just have been working on a slow pace for years within an organization. They are reasons for poor performance and slow growth.When they leave the organization the team becomes fast paced and the turnover time is decreased considerably. This means reduced cost .

DISADVANTAGES :

- **Decreased overall performance:**

The whole business process is affected when an employee leave the organization. It is even more risky when this happen all of a sudden. There is no time to train the new employee who is to take over the job and the whole team gets affected. It can directly be seen in an overall decrease of performance of the team. Sometimes this may even lead to drastic change in customer relationship. Customers connect with employees in an organization and those leaving all of a sudden may lead to doubts in customer's minds as well.

- **Daily task management:**

Sudden attrition may lead to difficulty in managing daily tasks. Even large organization struggle to manage their task when employees leave jobs, getting small information and managing daily tasks become difficult as they cannot be managed by small current team which is left behind.Organization generally have notice period to ensure there is a smooth transition but attrition states otherwise, employees who leave suddenly leads to unmanageable daily routines.

- **Increased cost:**

This has to be the highest disadvantage to a company when employees leave their jobs. There is increased cost associated with every level of the process – losing and paying the previous employee,hiring a new one, training cost for the new employee.Research shows that these costs are way more than the losses incurred in managing and missing out on work.

- **Lack of knowledgeable employees:**

This goes without saying when employees leave an organization they take with them the experience they have gained overtime.With organizations which has high attrition rate the average years of experience of employees is really low. This result in low performance, lack of loyalty and cluelessness on what company has been through.Older employees with their years of experience can take over critical matters which can never be trusted with these new employees.Even with employees who have experience are hired they may suffer at taking care of critical business matter as they are new to company's policies, culture and current employees.

- **Create a Negative image:**

It is not just that employees are looking for job, even organizations are on the outlook of qualified professionals. When any company has high attrition rate it negatively impact the brand of the organization.Recruiters' state that they find it difficult to map qualified candidates to the organization, as candidates opts out fearing the attrition rate. The reasons may vary but a negative image work against the organization.

8. CONCLUSION:

This paper presented the effect of voluntary attrition on organizations, and why predicting it is important. It further outlined various classification algorithms based on supervised learning to solve the prediction problem. The results of this research showed the superiority of the logistic regression classifier in terms of accuracy and predictive effectiveness, by means of the ROC curve. When used with its optimal configuration, it is a robust method that delivers accurate results in spite of the noise in the dataset, which is a major challenge for machine learning algorithms. The authors thus recommend the use of the logistic regression classifier for accurately predicting employee attrition in an organization, which enables HR to take necessary action for the retention of employees predicted to be at risk of leaving.

9. FUTURE SCOPE :

The dataset is a good representative of the general workforce in today's organizations. The good results from multiple classifiers justify that the features chosen are causes that contribute to voluntary attrition. Intuitively, data points that are close to each other are likely to have the same outcome of attrition. This is the basis for choosing the Logistic Regression algorithm in this paper. The Logistic regression classifier has good ROC-AUC and accuracy values. Instead of constructing a general model, it simply stores instances of the data and classifies by a majority vote of the classes of the nearest neighbors. Future work might include modifying the algorithm to weight neighbors so that nearer neighbors contribute more to the fit, rather than using uniform weights for all neighbors, and comparing results to the basic logistic model.