HR Employee Attrition Analysis

Project ID:2020MP092

Review -III

Group Members

RA1611003030436 Hansa Pradhan

RA1611003030430 Snehal Nivsarkar

RA1611003030410 Sujata Rai

RA1611003030408 Gursimran Singh

Supervised By:

Dr. Jabir Ali

Asst. Professor

Department of Computer Science & Engineering

Faculty of Engineering & Technology

SRM Institute of Science & Technology

Table of Contents I

1. Abstract
2. Literature Survey
3. Identification of Research Gap and Problem
4. Expected Impact on Academics/ Industry
5. Methodology of the Project Work
6. Detailed Design

Table of Contents II

7 Major Inputs (Infrastructure) Required

8 Results Obtained

9 References

Abstract

The project's main aim is to enable HR department to have a balanced approach towards the attrition process. Application does an analysis on the company's employee sentiments with the historical data and further predicts which of the current employees have the probability of leaving. It also tells us which factor is the most probable reason for the employee attrition across the company and focus on some required changes in the company for employee-friendly policy.

Literature Survey

- Jantan,Hamdan and Othman [2] have done research on Data Mining techniques for performance prediction of employees C4.5 decision tree
- Nagadevara, Srinivasan and Valk[6] tried to show relationship of withdrawal behaviors like lateness and absenteeism, job content, tenure and demographics on employee turnover .
- Hong, Wei and Chen [1]shows the feasibility of applying the Logit and Probit models to employee voluntary Turnover predictions.
- Marjorie Laura KaneSellers [3] explored various personal, as well as work variables impacting employee voluntary turnover

T

Identification of Research Gap and Problem

- Attrition has a very negative impact on an organization as it has various costs attached to it.
- In general the organisations do not have any structured approach for predicting employee leaving.
- The system uses the historical data of employees for better prediction on attrition but getting that data is not easy
- Most of the algorithms used for predictions give a little low accuracy
- Methods are being generically applied in order to control employee attrition. They are not specific to the needs of the employee.

Expected Impact on Academics/ Industry

- It can create value for the organization by providing insights into the employee behavior and being able to substantiate it with data.
- Retaining the best employees will be a great advantage for the company
- Predicting the attrition in time will help the company be prepared for the situation. It may even help them to stop the attrition by intervening in time.

Methodology of the Project Work

We will be using Random Forest machine learning algorithm for analysing the data because it has high accuracy.
It works best on small amount of data
A random forest is a tool for producing estimated probabilities in machine learning classification tasks.
One simply counts the fraction of trees in a forest that vote for a certain class.

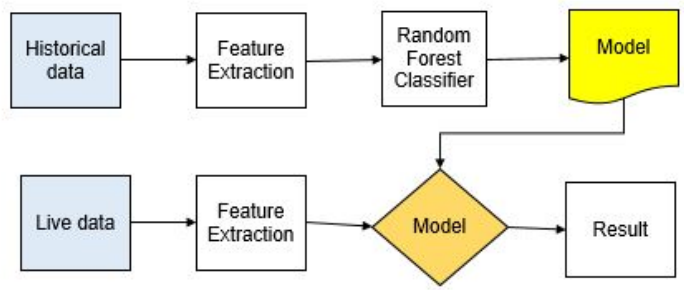

Figure: Random Forest Classifier model process

Detailed Design

- Analysis tool: This will help in predicting the attrition in the company's current data.
- Attrition List: This is the place where we display the people who have a high probability of attrition
- Visual Representation: All the data including the historical data is represented by in this attrition dashboard.
- Survey forms: There is also a sample survey form created as an example of how the data is collected from the employees regarding their details and how they feel about certain aspects of the company.
- Login Page: This is for the security of data so that only the HR people can acess the data
- Dataset : We are using a sample dataset of 1700 approx employees

Major Inputs (Infrastructure) Required

- Python Interpreter
- Any Integrated Development Environment(IDE)
- Any Browser
- All the python libraries which are required for the application
- Dataset

Implementation

The steps in the process flow shown in the Figure are explained as follows:

1. **Open application:** First, we open the application.
2. **Train the ML Model:** We need to train the ML model using Random Forest algorithm
3. **Attrition Categorization:** Random Forest will predict the class for the input given to be predicted
4. **View Attrition List:** This displays the list of people having high probability to leave.
5. **View Attrition Dashboard:** We can also see the visual representation of all the data.

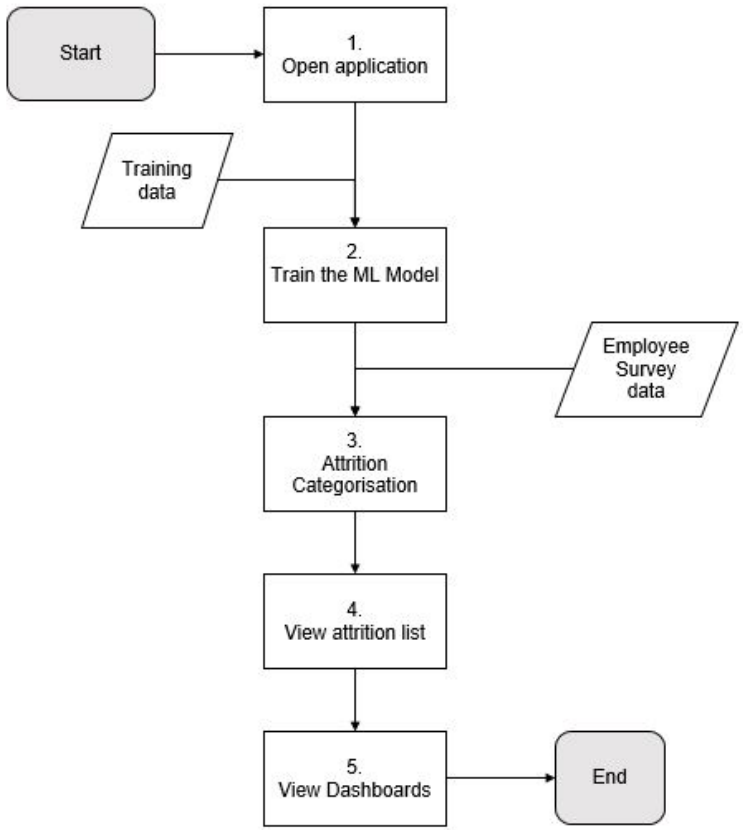

Figure: Process Flow

Results Obtained

Figure: login Page

Results Obtained 1

Figure: Home Page

Results Obtained 2

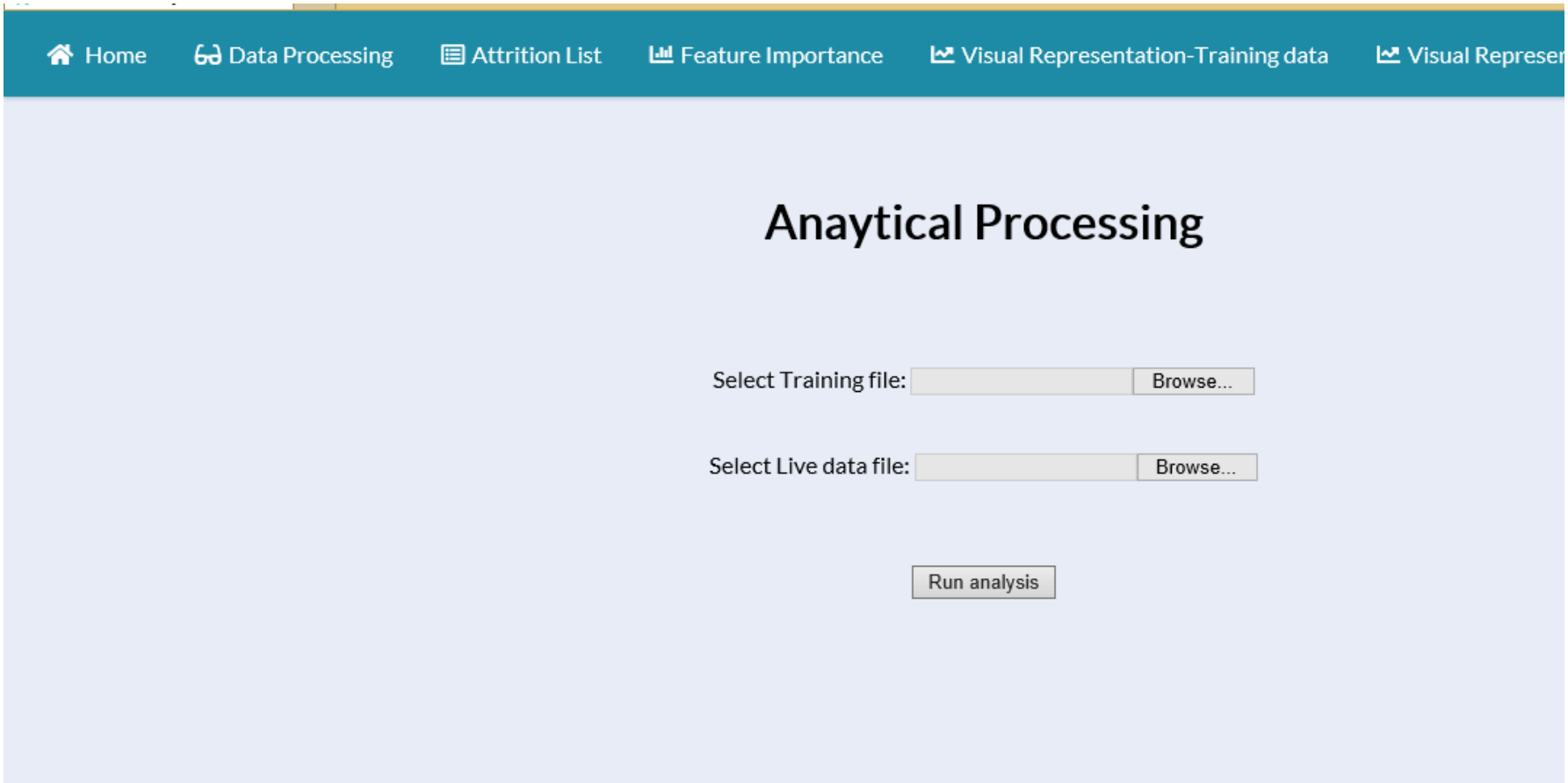

Figure: Analysis Tool

Results Obtained 3

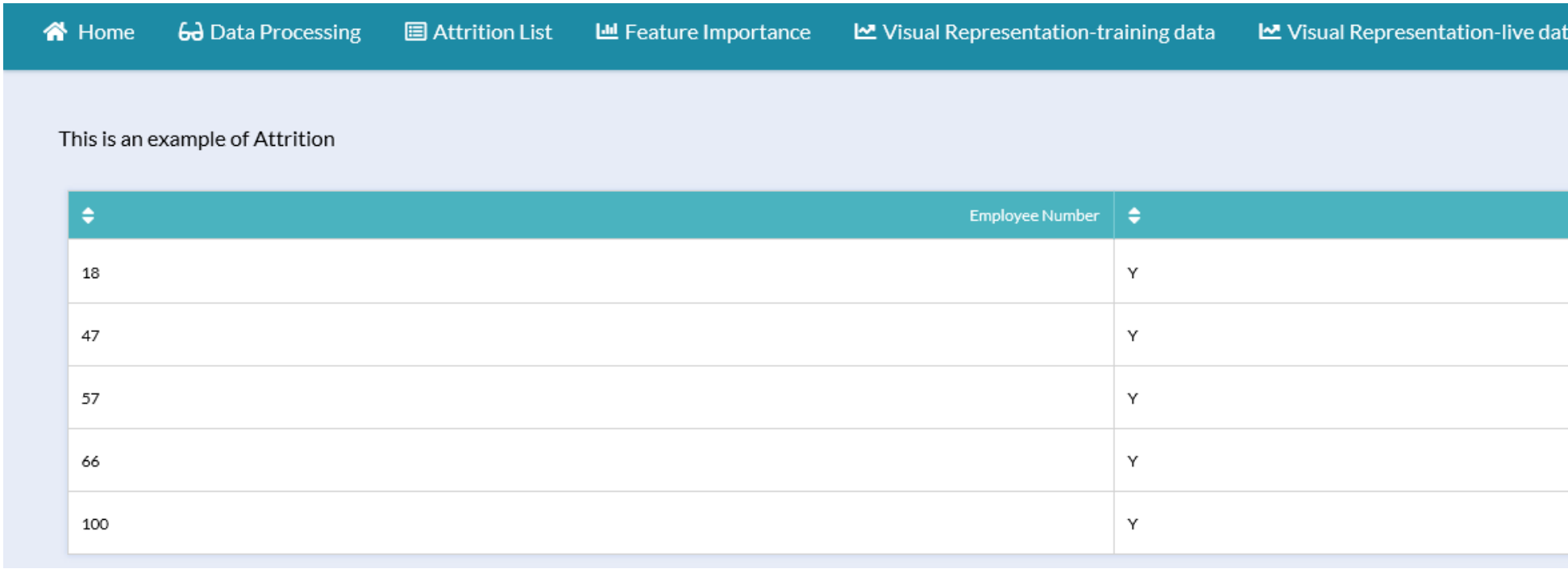

Figure: Attrition List

Results Obtained 4

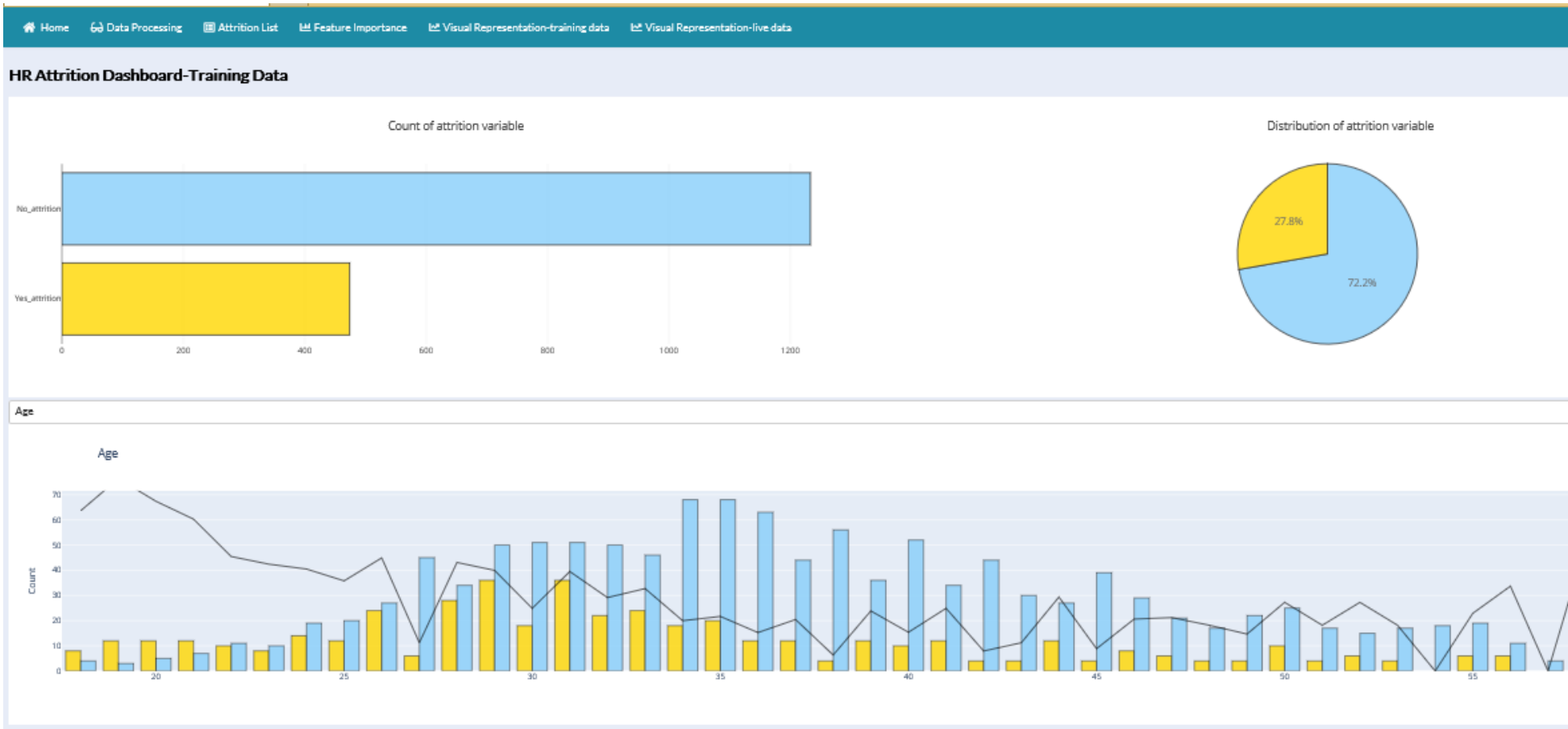

Figure: Attrition Dashboard

Results Obtained 5

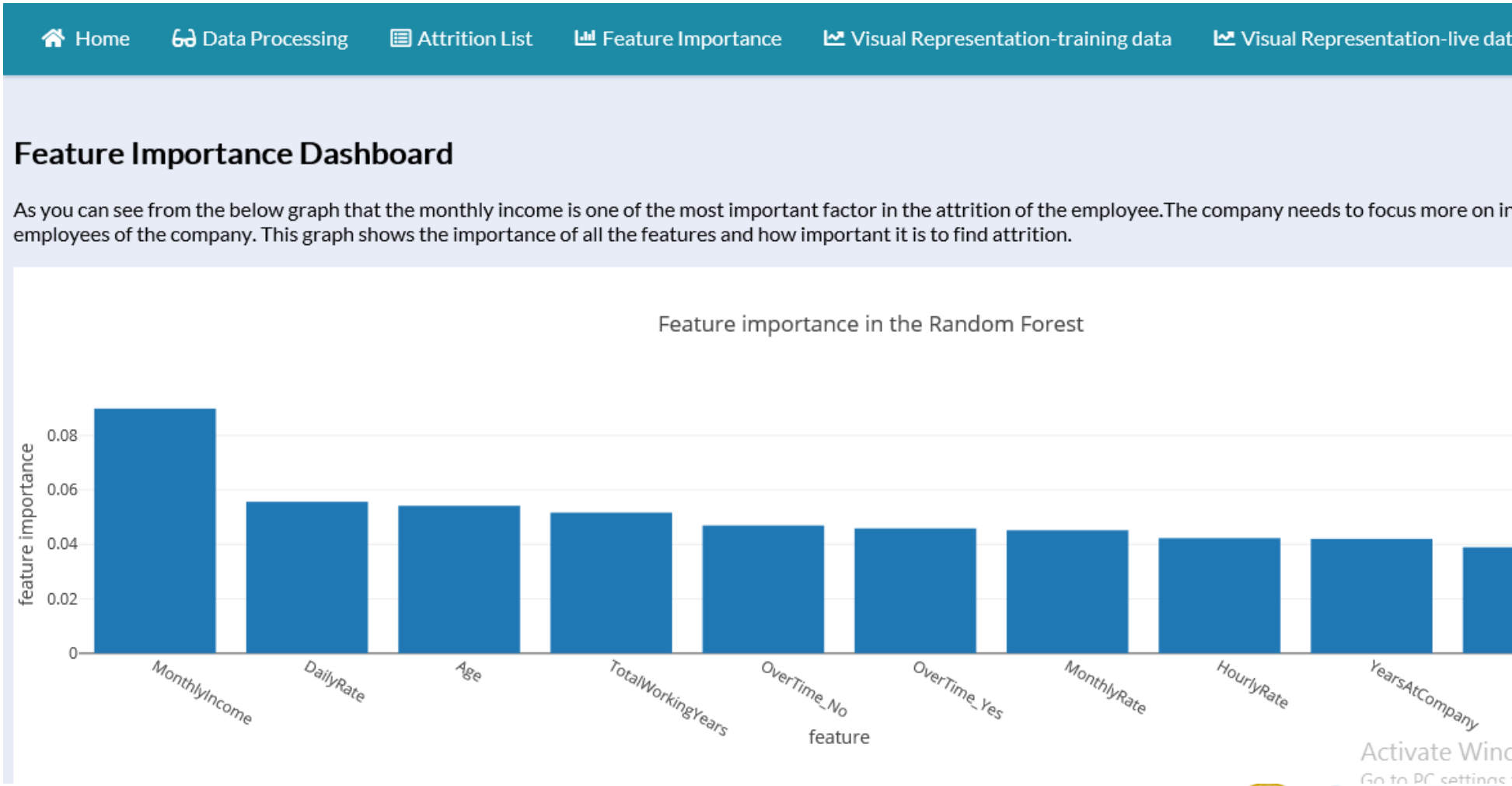

Figure: Feature Importance

References I


W.-C. Hong and R.-M. Chao.
A comparative test of two employee turnover.
International Journal of Management, 24(2):217, 2007.


H. Jantan, A. R. Hamdan, and Z. A. Othman.
Towards applying data mining techniques for talent management.
In *International Conference on Computer Engineering and Applications, IPCSIT*, volume 2, 2011.


M. L. Kane-Sellers.
Predictive models of employee voluntary turnover in a North American professional sales force using data-mining analysis.
Texas A&M University, 2007.

References II


D. Manivannan and M. Singhal.
Quasi-synchronous checkpointing: Models, characterization, and classification.
IEEE Transactions on Parallel and Distributed Systems, 10(7):703–713, 1999.


D. Manivannan and M. Singhal.
Asynchronous recovery without using vector timestamps.
Journal of Parallel and Distributed Computing, 62(12):1695–1728, 2002.


V. Nagadevara, V. Srinivasan, and R. Valk.
Establishing a link between employee turnover and withdrawal behaviours: Application of data mining techniques.
Research and Practice in Human Resource Management, 16(2):81–99, 2008.