# HR Employee Retention - Logistic Regression

To investigate the impact of various factors on employee retention in a company.

## Case Study

Employee retention is critical for maintaining productivity and reducing recruitment costs. High turnover rates can indicate underlying issues such as job dissatisfaction, lack of career advancement opportunities, or inadequate compensation. This study aims to identify key factors influencing employee retention and provide actionable insights to improve it.

## Dataset Description

[HR Analytics and Job Prediction Dataset](https://www.kaggle.com/datasets/mfaisalqureshi/hr-analytics-and-job-prediction)

Our data set consists of 14999 observations which include:

- Satisfaction level: Numeric value representing the level of satisfaction reported by each employee.

- Last evaluation: Numeric value representing the most recent performance evaluation score of each employee.

- Number project: Numeric value representing the number of projects each employee has been assigned.

- Average monthly hours: Numeric value representing the average number of hours worked per month by each employee.

- Time spend company: Numeric value representing the number of years each employee has spent in the company.

- Work accident: Binary variable indicating whether each employee has had a work accident (1 for yes, 0 for no).

- Left: Binary variable indicating whether each employee has left the company (1 for yes, 0 for no).

- Promotion last 5 years: Binary variable indicating whether each employee has been promoted in the last 5 years (1 for yes, 0 for no).

- Department: Categorical variable representing the department in which each employee works.

- Salary: Categorical variable representing the salary level of each employee (low, medium, or high).

## Libraries

- Pandas

- Matplotlib

- Seaborn

- Sklearn